Sources and Studies in the History of Mathematics and Physical Sciences

P. P. Divakaran

The Mathematics of India

Concepts, Methods, Connections

HINDUSTAN BOOK AGENCY

Springer

Sources and Studies in the History of Mathematics and Physical Sciences

Sources and Studies in the History of Mathematics and Physical Sciences was inaugurated as two series in 1975 with the publication in Studies of Otto Neugebauer's seminal three-volume History of Ancient Mathematical Astronomy, which remains the central history of the subject. This publication was followed the next year in Sources by Gerald Toomer's transcription, translation (from the Arabic), and commentary of Diocles on Burning Mirrors. The two series were eventually amalgamated under a single editorial board led originally by Martin Klein (d. 2009) and Gerald Toomer, respectively two of the foremost historians of modern and ancient physical science. The goal of the joint series, as of its two predecessors, is to publish probing histories and thorough editions of technical developments in mathematics and physics, broadly construed. Its scope covers all relevant work from pre-classical antiquity through the last century, ranging from Babylonian mathematics to the scientific correspondence of H. A. Lorentz. Books in this series will interest scholars in the history of mathematics and physics, mathematicians, physicists, engineers, and anyone who seeks to understand the historical underpinnings of the modern physical sciences.

More information about this series at http://www.springer.com/series/4142

P. P. Divakaran

The Mathematics of India

Concepts, Methods, Connections

Springer

HINDUSTAN
BOOK AGENCY

P. P. Divakaran (*emeritus*)
Kochi, Kerala, India

This work is a co-publication with Hindustan Book Agency, New Delhi, licensed for sale in all countries in electronic form, in print form only outside of India. Sold and distributed in print within India by Hindustan Book Agency, P-19 Green Park Extension, New Delhi 110016, India. ISBN: 978-93-86279-65-1 © Hindustan Book Agency 2018.

ISSN 2196-8810 ISSN 2196-8829 (electronic)
Sources and Studies in the History of Mathematics and Physical Sciences
ISBN 978-981-13-1773-6 ISBN 978-981-13-1774-3 (eBook)
https://doi.org/10.1007/978-981-13-1774-3

Library of Congress Control Number: 2018949905

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Preface

sadasajjñāsamudrāt samuddhṛtaṃ brahmaṇaḥ prasādena

sajjñānottamaratnaṃ mayā nimgnaṃ svamatināvā

(*Āryabhaṭīya, Golapāda* 49)

There is a good case to be made that the study of the traditional mathematical culture of India has now attained a level of maturity that allows us to begin thinking of going beyond just the description of its main achievements – theorems, constructions, algorithms, etc. – to attempts at a synthetic account of that tradition. Ideally, the scope of such a project should include not only the unifying mathematical ideas and techniques holding it all together, but also their place in the context of Indian history and Indian intellectual concerns. Its practical realisation will require many different skills and talents and is obviously a task for other times and other people. The more modest aim of the present book is to try and put together in a preliminary way what we already know from established sources, in the hope of drawing a first outline of its location within the universe of mathematics as well as, to a limited extent, in the Indian intellectual and cultural ethos.

The long and continuous history of the mathematical tradition of India was brought to an abrupt end in the 17th century. It was rediscovered – i.e., came to the notice of interested scholars outside the narrow circles of professional Indian astronomers (and astrologers) and mathematicians – some two centuries later. Many have contributed subsequently to the broadening and deepening of our knowledge and understanding of it. But areas of ignorance remained, of which the most serious and surprising was the astonishingly original and powerful mathematics created by Madhava and his long line of intellectual heirs in a small corner of Kerala, just before the arrival of European colonial powers on its shores. Recent work has begun to fill several such gaps and part of the aims of this book (as the footnotes will testify) is to give due weight to the results of the new scholarship. Two particular themes are worthy of special mention here. The first refers to what I have already invoked above: it is with a measure of disbelief that one realises that Madhava's name was largely unknown to modern historians of mathematics as late as thirty or forty years ago (and that the only translations of works by his followers, three in total so far, are not more than a decade old). So it is natural, for both historiographic

and purely mathematical reasons, that I should pay particular attention to the work of the school Madhava founded and the attendant historical and social circumstances.

The other area of darkness – perhaps the most resistant to illumination – on which current research is beginning to throw new light touches on the quotation at the head of this preface (an English translation will be found in Chapter 6.4), namely the long-held notion that mathematicians in India arrived at their insights in mysterious ways that are different from how 'everyone else' did it. Aryabhata, as anyone who has read him should have known, had no doubts: the "best of gems" that is true knowledge can be brought up from "the ocean of true and false knowledge" only by the exercise of one's intellect, by means of "the boat of my own intelligence". Others have said the same in less poetic language. The first vindications of this ringing endorsement of the rational mind, through the reconstruction of the actual proofs and demonstrations, are another gain from the current analytical studies. I have thought it essential therefore to provide complete – though not always elaborate – proofs of most of the significant results in modern notation and terminology, either as direct transcriptions or by connecting together partial information available in different texts. The hope is that we can begin to chip away, little by little, at the encrusted layers of the mythology – and, occasionally, ill-informed bias – that still has some currency: that Indian mathematics was, in some strange and exotic way, not quite of the mainstream. If this book has an overriding message, it is that, conceptually and technically, the indigenous mathematics of India is in essence the same as of other mathematically advanced cultures – how can it be otherwise? – and that its quality is to be adjudged by the same criteria.

In navigating the ocean of true and false historical knowledge – an issue that has a special relevance in today's India – I have had (unlike, it goes without saying, Aryabhata) the generous help of many mentors, colleagues and friends. But before all of them must come the late K. V. Sarma whose work, more than anyone else's, helped place Madhava among the mathematical greats of all time, and not only of India: a fortuitous meeting with him while he was translating *Yuktibhāṣā* was the spark that first lit my interest in the subject. Among the many others who sustained it by sharing their expertise or in other ways are Ronojoy Adhikari, Kamaleswar Bhattacharya, Chandrashekhar Cowsik, Naresh Dadhich, S. G. Dani, Pierre-Sylvain Filliozat, Kavita Gangal, Ranjeev Misra, Vidyanand Nanjundaiah, Roddam Narasimha, T. Padmanabhan, A. J. Parameswaran, François Patte, Sitabhra Sinha, N. K. Sundareswaran, $\cdots$. A special word of appreciation and gratitude is due to Bhagyashree Bavare, my collaborator and consultant in matters Sanskritic, as well as to Samir Bose and N. G. (Desh) Deshpande who read the book in draft form and made many valuable suggestions. Above all, this book and I owe much to four old friends: S. M. (Kumar) Chitre for his faith and encouragement at a time when they were needed; David Mumford with whom I have had many stimulating exchanges about the content and context of the

mathematics of India; M. S. Narasimhan whose insights into what one may call the universal mathematical mind, freely shared over a long time now, have contributed greatly to my understanding of its Indian manifestation; and the late Frits Staal whose broad vision of India's intellectual heritage has strongly influenced my own thinking. Institutional support and hospitality at various times in the preparation and writing of the book came from the Institute of Mathematical Sciences in Chennai, the DAE/MU Centre of Excellence in the Basic Sciences in Mumbai, the Inter-University Centre for Astronomy and Astrophysics in Pune and the National Centre for Biological Sciences of the Tata Institute of Fundamental Research in Bengaluru. Finally, I am happy to acknowledge here my gratitude to the Homi Bhabha Fellowships Council for choosing me several years ago for a Senior Award, which event was the trigger that set off the idea of putting it all down on paper.

P.P. Divakaran
February 2018

Contents

Introduction

0.1 Three Key Periods

This book has four parts. The first three correspond to chronological divisions of the long history of mathematical work done in India while the last part is a bringing together of the unifying ideas and techniques that run through that history. Two of our three historical periods are obligatory choices: the very beginning of mathematical thought in the Vedic age (earlier than ca. 500 BCE) as far as we can tell from available records, and the last phase (ending ca. 1600 CE) which is, naturally, much better documented. It is a remarkable fact that the end did not come as a slow fading away but, on the contrary, was relatively abrupt following a period of blazing glory; so it is doubly worthy of a detailed study. Between the beginning and the end is a middle or classical period spanning roughly the two centuries starting around 450 CE which, as it happens, is also of pivotal importance. This is the period of Aryabhata (Āryabhaṭa) and Brahmagupta, one in which the Vedic mathematical legacy was radically transformed.

As in all ancient mathematical cultures, so in India: mathematics began with learning how to enumerate discrete sets (counting) and how to create simple planar figures (drawing). The former took the form of decimal enumeration and led in due course to arithmetic and the latter to a fairly elaborate geometry, primarily of the circle and of figures associated with the circle. Historically, these first awakenings of the mathematical spirit span the Vedic age beginning, say, around 1,200 BCE if we go by textual evidence alone. There are, however, intriguing glimpses of an earlier familiarity with some of the Vedic geometry in the archaeological remains of the Indus Valley (or Harappan) civilisation, going back to the middle of the 3rd millennium BCE or a little earlier. What is certain is that the basic principles of both decimal arithmetic and geometry were pretty well in place at least by the time of the composition of the earliest of the architectural manuals called the *Śulbasūtra*, ca. 800 BCE, possibly somewhat before that. We do not have much information on the further development of geometry for a long time after, in fact until the classical period. But the consequences of the idea of measuring a number (I use 'number' without qualification to mean a positive integer), no matter how large, by means of the base 10, in other words decimal place-value enumeration, continued to be

vigorously explored until the beginning of the common era and beyond. Historians are of the view that during this period, the Vedic people were in the process of settling the Gangetic plain, setting out from the upper Indus region and working their way gradually towards the east.

The classical period, which got its chief impetus from astronomical investigations, is characterised by several 'firsts'. To begin with, astronomy itself became a science: the motion of heavenly bodies came to be studied as the change in the geometry defined by them, considered as points in space, with the passage of time. The required geometry was seen to need a reformulation and refinement of that inherited from Vedic times. Simultaneously, the reconciliation of time measurements made using the lunar and solar periods led to problems requiring the solution of algebraic equations in integers (and of course, first, a proper algebraic formulation of such equations and their properties). Circle geometry became greatly sophisticated, going beyond the needs of astronomy. Also from this phase we have the first texts concerning themselves exclusively with mathematics and astronomy which are, moreover, attributable to historically identifiable authors: Aryabhata first and foremost, followed a little over a century later by Brahmagupta and Bhaskara (Bhāskara) I. For those who know even a little about the story of mathematics in India, the mention of these names is enough to indicate the significance of this period for everything that came later. Indeed it is not an exaggeration to say that all of the subsequent work in astronomy and related mathematics in India is an elaboration of the concepts and techniques first formulated during this time.

The geographical locus of mathematical activity during this middle phase was fairly widely distributed within northern India; Aryabhata did his work in Magadha, part of modern Bihar, while Brahmagupta is thought to have lived in Malava (Mālava), the region around Ujjain (Ujjayini).

My third chronological division corresponds also to a geographical division and it is defined entirely by the achievements of Madhava (Mādhava). Madhava and a long line of his disciples lived and worked in the lower basin of the river Nila (Niḷā) in central Kerala and they remained productive for two centuries, ca. 1,400 - 1,600 CE. It is by far the best documented phase of Indian mathematical history. Despite that, it has taken a long time for modern scholars to go from relative ignorance to puzzled admiration to an informed appreciation of the brilliance and originality of the achievements of this last phase.

The central concern of the Nila school[1] was the conceptualisation and execution of a method of dealing exactly with geometrical relationships in nonlinear problems, specifically problems of rectification and quadrature involving circles

[1]The name is meant to convey not only the thematic coherence of the mathematics it produced but also its localisation in a cluster of villages all within easy walking distance of each other. Other designations such as the Kerala school and the Aryabhata school are used by some authors. The first ignores the fact that there was a shortlived but vibrant astronomical research centre elsewhere in Kerala in the 9th century CE. As for invoking Aryabhata's name to describe their work, almost all of Indian mathematics from the 6th century onwards can with justice be called the Aryabhata school.

and spheres, by resorting to a process of 'infinitesimalisation' through division by unboundedly large numbers. In other words, the fundamental achievement of Madhava and the Nila school was the invention of calculus for trigonometric functions (as well as, along the way, for polynomials, rational functions and power series). The roots of the Nila work go directly back to Aryabhata, skipping over much of what happened in between, and, in a very real sense, it represents the coming together of the two main strands of the earlier tradition, geometry (of the circle and associated linear figures) and based (decimal) numbers. In the process, it sharpened and deepened algebraic methods traceable in a formal sense to Brahmagupta, defined and made respectable entirely new types of mathematical objects (infinite series) and introduced methods of proof (mathematical induction) unknown till then.

One would expect a time of such effervescence and creativity to be followed by a period of consolidation and steady progress. But that is not what happened. The Nila school marks the final episode in the long and essentially autonomous progression of mathematical thought in India. Little of any value or novelty emerged after the 16th century in Kerala or indeed in India as a whole. Quite a few reasons have been put forward for the decline – it is not every century that produces a Madhava – but it cannot be doubted that the immediate trigger was the arrival of Portuguese colonialism on the shores of Kerala.

The choice of the three epochs that we are going to concentrate on is thus not random. Historically and in terms of mathematical impact, they represent the highest of high points in a long evolution. What of the gaps – each, coincidentally, of about a thousand years, give or take a century or two – that separate them? By and large, they served as periods of steady if unspectacular advances. There were of course very many fine mathematicians and astronomers, especially between the middle and the final phases, who do not fit into our overneat division, the most outstanding being the versatile and prolific Bhaskara II. Their contributions were handsomely acknowledged by those who came later and will figure in our account – Bhaskara II in particular had an enormous influence and his *Līlāvatī* is probably the most popular mathematical text ever written in India. The hope is that what emerges at the end is a well-rounded portrait of mathematics within the boundaries of cultural India – which rarely coincided with the various political Indias of kings and empires – and over its recorded history. The informed reader will judge how well the hope is realised.

In the short concluding part I have tried to flesh out the sense of continuity and unity in the mathematical culture of India, already visible to the enquiring eye in the first three parts, by gathering together some of the common strands of thought and method that run through it all. They relate to foundational issues as well as to the specific means employed to get specific results - in fact the two aspects are so inextricably joined that they appear to be different expressions of an all-pervasive common mode of thought. The prime example of this unvarying 'mathematical DNA' is from geometry which is almost always reduced, from the *Śulbasūtra* down to the Nila texts, to the application of two principles: the

Pythagorean property of general (i.e., not restricted to integral or rational) right triangles and the proportionality of the sides of similar (generally, right) triangles. Similarly, the recursive principle underpinning the construction of decimal place-value numbers and the early realisation that they are unbounded have big roles to play in the treatment, 2,500 years later, of power series in the Nila work. Even the inductive proofs which make their first appearance in this last phase are introduced from a recursive standpoint, as distinct from the formal-logical foundation favoured later in Europe.

Just as conspicuous as these commonalities is the absence of concepts and methods which the modern mathematically literate reader accepts unquestioningly. For instance, though the operation of division and the notion of divisibility are present from very early times – the place-value construction of numbers is based, after all, on the iterated application of the division algorithm – there appears to have been no special significance attached to prime numbers, and it is not clear why. Less puzzling is the avoidance of the powerful and ubiquitous method of proof based on the rule of the excluded middle, proof by contradiction. The logical basis of reasoning by contradiction was considered by philosophers from a very early time, pronounced unsound and rejected as a means for the validation of knowledge. A belief used to be prevalent in some mathematical-historical circles (and is occasionally still expressed) that Indian mathematicians did not have proofs for many of their assertions or, worse, that a conception of a logical proof itself was absent. It will be seen that this is a myth; only, the logic is somewhat different from what is axiomatic in the European approach, that of a progression from axiom to definition to theorem using, at every stage, equally axiomatic rules of reasoning.

0.2 Sources

The term ‘India’ or, occasionally, ‘cultural India’ as used in this book encompasses of course political India at the time of independence but also covers what is called the Indian subcontinent, extending well into Afghanistan in the northwest. Historically and culturally it also included large parts of southeast Asia which, as will be seen, made at least one significant contribution to the reconstruction of the whole story.

The recorded evidence on which the reconstruction must be based comprises any surviving material, concrete or (more or less) abstract, from which we can draw reasonably secure conclusions about those preoccupations of the past inhabitants of India that have a clear mathematical content. The vast majority of such records are texts dealing with mathematics, generally in parallel with applications to some practical science which, in later times, was almost invariably astronomy. But there are other sources of information as well which supplement the texts and make up for their shortfalls. An extreme instance is the earliest historical period, the Indus Valley period, which has left behind a splendid collection of ruins but whose writing has not yet been deciphered.

The only recourse we have then is to try and 'read' civilisational artefacts other than texts, abstractions such as townscapes and street plans, floor plans and elevations of buildings, geometric decorations on seals, pottery, etc., as well as material objects like weights and other measures. While little can be said with absolute certainty, surprisingly good information of a general mathematical nature can be plausibly extracted from Indus Valley material of various sorts.

Textual sources can be divided into two types to begin with, those whose mathematical content is scattered and incidental and those devoted wholly to expositions of mathematics (and its applications). Most of the early texts of either type were composed orally, memorised and, in the early part of their existence, transmitted orally before being transcribed on perishable material. But before turning to these two types, I should mention a different kind of written evidence, inscribed permanently (very permanently) on stone monuments, coins, copper plates recording grants etc., the most famous being the first symbolic representations of zero as an approximate circle carved on stone (southeast Asia and Gwalior in India). They are comparatively late, the earliest numerical inscriptions (outside the Indus Valley seals) being from the centuries around the turn of the common era, showing numerals in the Brahmi (Brāhmī) script. The late occurrence of written numerals is consistent with the primacy of number names rather than number symbols in India, a topic which will occupy us later. After they came into common use, Brahmi numerals seem to have gone through a complicated evolution before settling down to the currently standard Arabic numerals as well as various minor variants in Indian languages. They are well documented in several studies, notably in the book of Datta and Singh ([DS]).

The texts I have characterised as incidentally mathematical are of paramount importance as they are, with the exception of the *Śulbasūtra*, our only primary source of knowledge about mathematics in the Vedic period, in particular about the genesis and rapid development of decimal place-value enumeration and its use in elementary arithmetic. They comprise, first, the earliest Indian literary productions, namely two of the original Vedic (*saṃhitā*) texts, the *Ṛgveda* and the *Yajurveda*, the latter more particularly in the earlier of its two recensions, named *Taittirīya Saṃhitā*. Some useful information also comes from a few of the *Brāḥmaṇa*s and *Upaniṣad*s which are, by and large, exegetic elaborations of Vedic ritual and philosophy. Scholars date the gathering together of the *Ṛgveda* into ten Books or 'circles' to around 1200 - 1100 BCE but also think that the individual poems or hymns were composed in the centuries leading to that period from the time of the appearace of the Vedic people in northwest India. The *Taittirīya Saṃhitā*, which is mostly a text describing details of ritual practices, is probably from only a little later and is full of various regular sequences of numbers, including lists of names of powers of 10. It would seem from the way these lists are presented that the first intimations of the unboundedness of numbers were already in the air. The fascination with the potential infinitude of numbers continued for a long time – perhaps not

so long considering that the quarry being chased was unattainable – leading to the naming of sequential powers of 10 extending to mind-boggling magnitudes. Evidence for this comes from religious texts (Vedic/Hindu as well as Jaina and Buddhist) and from other sources like the great epics *Rāmāyaṇa* and *Mahābhārata*, composed in the form we know them today probably just before the common era.

The value of literary and religious texts as repositories of mathematical information is greatly enhanced by the surprising – for a people who turned everything they knew into literary compositions – absence of any work that lays out the principles governing decimal enumeration and decimal arithmetic. That is not the case when it comes to early geometry. The *Śulbasūtras* are traditionally said to have existed in a number of versions of which four appear to have survived more or less intact. They are thought to be different recensions of material drawn from a common store of knowledge, made at different times (and perhaps in different places in the Gangetic basin) ranging over ca. 800 - 400 BCE. Being handbooks for the construction of Vedic ritual altars, their purpose is primarily architectural but the precise and quantitative treatment of floor plans having different regular shapes turns them effectively into workbooks of plane geometry. In this respect the *Śulbasūtra*s stand apart from all the other sources we have for mathematics in the early period.

The middle or classical period brought about a radical and permanent change in the nature of mathematical enquiry and, with it, in the character of the texts. Astronomy became the driving force of mathematics; in fact the two sciences became inseparably linked. Going by the texts, all astronomers were also mathematicians and most mathematicians were astronomers, a trend that survived right until the last phase. It is tempting to say that most of them would have thought of mathematics as the prerequisite for doing astronomy rather than of astronomy as applied mathematics. (From now on it will be understood that the word 'mathematician' as used in this book will refer generally, with one or two obvious exceptions, to an astronomer-mathematician.)

As for their texts, I have already noted that the classical period saw the appearance of what we may call monographs for the first time (the *Śulbasūtra*s partially excepted). Their production became increasingly frequent and one has the sense that this period also marks the emergence of a professional mathematical community.

The texts themselves can be divided loosely into two broad categories: original works having some new mathematics to expound and commentaries (*bhāṣya* or *vyākhyā*; there are also other less common names) on them (and an occasional commentary on a commentary). The number of commentaries that a particular work generated seems to have been an increasing function both of its significance (and, hence, of the esteem and reverence in which its author was held) and of its brevity (and, hence, of its difficulty). The prime example is of course *Āryabhaṭīya*, highly esteemed and very short, with about half a dozen known commentators extending from Bhaskara I (early 7th century) to Nilakantha (Nīlakaṇṭha) (early 16th century) trying their hand at making its

cryptic verses more accessible; indeed, it is only a mild exaggeration to characterise most of the work during this time as constituting a working out – a multi-author *mahābhāṣya* so to say – of Aryabhata. Other influential examples are *Brāhmasphuṭasiddhānta* of Brahmagupta, some of the writings of Bhaskara II and, from the late phase, Nilakantha's *Tantrasaṃgraha* which inspired two very important self-described commentaries on it in less than 50 years. Naturally, the commentaries also tend to be longer; and they are often in prose, more frequently than the source texts to which they are anchored.

Nonetheless, the division into these two categories is not watertight. Firstly, every text starts with a résumé of the knowledge already acquired, whose purpose was to serve as the platform on which new mathematics will be built – even *Āryabhaṭīya*, possibly the shortest mathematical treatise in history, begins with an evocation of decimal enumeration by recalling the names of powers of 10 as well as some elementary geometry. For the historian this brings in two advantages: filtering out (the very infrequent) patent errors from earlier work while at the same time giving a glimpse of how later mathematicians understood (or, occasionally, misunderstood) the material they were building on. We shall meet instances of both in the course of this book.

Conversely, many commentaries – so described in their titles or opening passages – not only provide clarifications and interpretations of the original work, but also put forward genuinely new ideas and insights. A very good example is Nilakantha's *Āryabhaṭīyabhāṣya* which subjects Aryabhata's text to a searching analysis and reinterpretation in the light of the new results of Madhava and the Nila school and, in the process, turns some of the imprecisely worded aphorisms in the original into precise mathematical statements which Aryabhata himself may or may not have had in mind.

The generally accepted explanation for most of these differences of presentation is that the original texts were intended to be memorised as the core of a guru's teaching, to be supplemented by face-to-face explanations and demonstrations. In any case, no modern reading of an ancient mathematical classic can be considered complete without attention being paid to the more important commentaries that came later, even in the rare cases where they may have been misled.

The language of Indian mathematics is, overwhelmingly, Sanskrit. From Vedic times until the very end, from the northwest to the extreme south and at all places in between, all the main texts – with one very significant exception – were composed and communicated in the variant of Sanskrit prevalent at the particular time, without much geographical variation. During the first period when Vedic Sanskrit was the natural language of literate people and the second period when classical Sanskrit was the language of learning in north India, this is to be expected. As mainstream mathematics spread further south to regions with a strong local literary and linguistic life of their own, the habit of doing traditional sciences, the *śāstra*, in Sanskrit survived. An explanation for this conservatism may be that Brahmins, who were the people of learning and who began migrating to south India in large numbers in the second half of the

first millennium CE, remained protective of their inheritance for a long time. Apart from some historical evidence that this was so, quite strong in the case of Brahmins in Kerala, it is a fact that the style of formal Sanskrit mathematical presentation hardly changed from Aryabhata to Nilakantha.

The significant exception mentioned above is the remarkable treatise *Yuktibhāṣā*, written in Kerala around 1525 CE by a "venerable twice-born" (Brahmin) named Jyesthadeva (Jyeṣṭhadeva). It is remarkable for many reasons as we shall see – a large part of Part 3 of the present book depends critically on its reading – but what is to be noted here is that it is not in Sanskrit but in Malayalam, the language of Kerala. And it is not as though it marked the beginning of a trend; several works were composed in Kerala during the following century in chaste classical Sanskrit. We have no idea why Jyeshthadeva decided to popularise, if that was the intention, a science which until then had remained the preserve of those who were proficient in Sanskrit.

Another thing we do not know is when the texts began to be written down. There is little doubt that material from the Vedic age was communicated orally for a long time. Even in the classical age, by which time writing in various scripts had become fairly standard, oral teaching and learning still retained its preeminence. Nevertheless it is a tacit assumption in Indological studies, very likely correct, that most material got written down sooner or later. (I will use terms like 'writings' and 'books' to refer to work which may have begun their life orally, as I have used 'literate' to mean also orally literate beyond the needs of basic functional communication). In any case, what was actually written down was on perishable material like palm leaves and manuscripts had to be copied and recopied periodically for survival. Consistent with the history of the influx of Brahmins into parts of south India, most mathematical manuscripts originating in northern India (of the classical age and later) were unearthed in Kerala, the last refuge of the sciences in colonial India, often from the private libraries of Brahmin families. At the same time, the impact of Sanskrit on the local language gave rise to modern Malayalam, written in a rapidly evolving script of its own. At some point in this process of sanskritisation, Sanskrit itself came to be written in the local script, with the end result that the vast majority of surviving Sanskrit manuscripts in Kerala – i.e., a large proportion of *all* surviving manuscripts – are in Malayalam script. It will obviously be absurd to conclude from this, as some writers still do, that Aryabhata for example was born in Kerala.

In the light of all this, it is natural to wonder whether we have in hand today the essential primary sources necessary for the delineation of a satisfactory picture of the totality of Indian mathematical culture, in other words whether some critically important manuscript may not still be lying unsuspected in the recesses of some old family library or lost permanently to the ravages of time. We cannot of course be sure but, here again, the persistence of knowledge acquired earlier across generations of later writings provides a guarantee, I think, that nothing truly significant has disappeared for good. There was a time in the first half of the 19th century when scholars had access to Brahmagupta's

writings with their numerous and hostile criticisms of Aryabhatan doctrine, but no first-hand knowledge of the *Āryabhaṭīya* itself. The world was saved the need to guess what the *Āryabhaṭīya* really contained by its publication in 1874. The nearest to finding oneself in a similar predicament again is the absence of any known mathematical text – a fragmentary astronomical work of little mathematical interest apart – attributable to Madhava. We can always hope that such a manuscript will one day come miraculously to light but that is unlikely to change our understanding of his mathematcs; the extensive writings of those who followed him in the Nila school already give us a very coherent idea of what it might contain.

There are also issues on which the modern reader would prefer to get a little more help than is available in the texts. The first concerns diagrams. Perhaps because drawing with a heavy iron stylus on dried palm leaf is not easy, they are absent from manuscripts except for an occasional rudimentary sketch. This can be a handicap when trying to follow a complex line of geometric reasoning (examples are Brahmagupta's theorems on cyclic quadrilaterals and the infinitesimal geometry of the Nila school). What there is instead are descriptions of the necessary figures, which are generally adequate. In face-to-face instruction, diagrams were presumably drawn on sand covering a board or the floor, a practice which was prevalent in Kerala until as late as the middle of the 20th century.

The other striking absence is that of a symbolic notation for mathematical objects and operations on them even where long and intricate and fundamentally algebraic reasoning is involved, as in Brahmagupta's treatment of the quadratic Diophantine equation (later called Pell's equation in Europe) or in the derivation of trigonometric infinite series by the Nila school. It is possible that this is a legacy of the oral tradition, but it is a fact, nevertheless, that abstract representations of linguistic objects and sets of such objects were already employed by Panini (Pāṇini) (ca. 5th century BCE) in his treatise on grammar. In mathematics, the use of syllables to stand for unknown numerical quantities was initiated by Brahmagupta and given much prominence by Bhaskara II in his monograph on algebra (*Bījagaṇita*). But one will search in vain for a symbolically expressed equation. In consequence, a considerable part of the effort of reading a text has to go into turning narrative descriptions into the present day mathematical language of symbols and symbolic statements of relations among them. Far more seriously, the disdain for a symbolic presentation may be a reflection of a certain methodological position which in turn led to a reluctance to see mathematical truth in full generality and to obvious paths not taken. We will return to this point in somewhat greater detail in Part IV.

On the other hand, the absence of detailed proofs in many of the texts, whether originals or commentaries on them, is not a serious matter. With very few exceptions, proofs of stated results, either complete or detailed enough for an informed reader, can be found in some text or the other. We have come to see recently that even in accounts in which the logical development towards

a final statement ('theorem') is supposed not given, a proper understanding of the fine distinctions in the terminology employed may in itself be a good guide to the reconstruction of the logic. More generally, the way a proof – the commonly used word is *upapatti* or, in later times, *yukti*, 'reasoned justification' – is put together and what it assumes as self-evident rules of reasoning are quite different from the modern approach to such questions. Logicians throughout history have wrestled with the issue of what constitutes a good criterion for the validation of new knowledge. In this book I use the word 'proof' without further qualification in the sense accepted by Indian mathematicians – we shall encounter numerous examples below – even though that may not meet with the approval of axiomatists ancient or modern; Indians did not believe in axioms. It would seem that just as Indian mathematical authors have their own particular style of presenting their knowledge, so does it require a particular way of reading to absorb it in full measure and to be convinced by it.

This short general survey of the sources will be incomplete without a look at one written work to which much of the general discusion above applies with a high degree of reservation. The Bakhshali manuscript as it is known was unearthed (literally) in 1882 in what is now Pakistan, within or close to the borders of ancient Gandhara. It is written on birch bark, a popular medium of writing from at least the beginning of the common era in the regions of north India where the Himalayan birch grows. The fraction – how small a fraction remains unknown – of the complete text that has survived is in such a fragile condition as to preclude scientific methods of dating (in the judgement of its keepers). Whether it is a copy or a recopy of a preexisting work or the original is also not known. In the latter case, the lack of a scientifically determined date is especially unfortunate because its contents are in many respects singularly novel. As it is, based on its mathematics, its language (a local version or dialect of Sanskrit) and script (also a local variant of the Brahmi/Devanagari family), not to mention personal preferences, people have dated it anywhere from the 2nd century BCE to the 13th century CE. Even if we discount the extremes of this quite impossibly wide range and date its substance between, say, the 4th and the 8th centuries CE, the historical conclusions we can draw from it will depend dramatically on precisely when it was composed since these centuries encompass the transition to the classical age and a very rapid evolution within that period. In particular, an early, say pre-Aryabhatan, date – I will in fact argue later for just such a position – will force a radical revision of our present understanding of the evolution of arithmetical and algebraic ideas in the gap between our first two periods.

A quick look at the contents of Bakhshali shows us how and why. It is almost unique among the standard texts in utilising, at several levels, a fairly evolved symbolic notation. First, numbers are written in symbols (as distinct from their names) in decimal place-value notation using a variant of Brahmi numbers. (The earliest Brahmi numbers on inscriptions etc. that I mentioned earlier did not follow a strict place-value notation). Then, arithmetical operations were also represented abstractly (often in addition to a narrative descrip-

tion) and that encompassed both negative numbers and fractions as well as schemata for equations. As cannot be avoided in a written symbolic notation, there is a symbol (a dot) for zero. And finally, there are algebraic equations with a symbol representing the unknown. These features signalling an evolved arithmetical culture would indicate a relatively late date for Bakhshali if we insist on dated, independent corroboration, e.g., on Brahmagupta for the algebra of positives and negatives or for the use of syllables for unknowns. But if Bakhshali itself could be reliably dated to be early within its plausible range of dates, that would push back the time of the first signals of an algebraic mode of thinking by a few centuries and thereby give us a measure of the progress made in the interregnum between the end of the Vedic phase and the beginning of the classical phase. Such a chronology is not out of the question; the *Śulbasūtra*s do contain some arithmetic with fractions and subtractions (negatives?). Notational devices like those in Bakhshali resurface in one or two later texts – the great classical works avoid them – lending support to the view that they were in continual use behind the scenes, in traditional teaching, in the classroom as it were.

The decline in creative mathematical activity in Kerala (and India as a whole) that followed the arrival of the Portuguese caused a corresponding decrease in the number of new texts as well. It is true that the 16th century saw the production of some masterpieces, *Yuktibhāṣā* among them, but, with one or two exceptions, there was not a great deal that was original, even in terms of interpretation, in the few serious books that were written after the 16th century. What did happen instead, beginning in the late 18th century, was a slow rediscovery of traditional Indian learning, including mathematics and astronomy. Naturally, the pioneers in this endeavour were British Sanskritists, following in the footsteps of the first and the finest of them, William Jones of Calcutta. Towards the end of the 18th century and in the first few decades of the 19th came the first translations, notices, commentaries and lectures (the translation of a part of *Sūryasiddhānta* by Samuel Davis and John Playfair's essays based on it, Colebrooke on Brahmagupta and Bhaskara II, Whish on some Nila texts, and several others). The second half of the 19th century saw a gradual acceleration of this activity in a number of European countries and even in America, bringing several key texts to the notice of scholars, *Śulbasūtra* and *Āryabhaṭīya* among others. Thus came about the slow and hesitant process by which the world came to recognise the existence of an Indian culture of mathematical enquiry that was in large measure self-motivated and self-sufficient.

While a residual knowledge of traditional texts and their utilisation in preparing almanacs survived through colonial times, the reawakening of a critical and informed interest in their own true mathematical heritage among Indian scholars came only towards the end of the 19th century. It has grown vigorously since. Given also the continuing activity in other parts of the world in bringing out these riches that long lay hidden, this seems a propitious time to attempt an account of mathematics in India – the mathematics of India – that not only explains its high achievements but also pays attention to its own special char-

acteristics, its Indian identity as it were. It goes without saying that any such attempt must rely on the work of the fine line of scholars, extending over more than two centuries now, who have brought their linguistic and mathematical skills to the elucidation of what, for many of them, is an alien way of doing and presenting mathematics. The informed reader will find many instances in this book of my indebtedness to the insights of these secondary sources. It is astonishing to realise, nevertheless, that there have only been three books ever that have taken on the task of presenting a wide-ranging and mathematically oriented account of the subject. Of these, the otherwise encyclopaedic two-volume book of Datta and Singh ([DS]) published in the 1930s omits geometry (the promised third volume never came out except, much later, as a posthumous journal article), but Sarasvati Amma's comprehensive and magisterial book ([SA]) – faithful and lucid at the same time – more than compensates for the omission. The very recent book of Kim Plofker ([Pl]) covers most of the geography and history of Indian mathematics and has an extensive bibliography of current research. All three books are essential reading.

0.3 Methodology

The present book differs from those referred to above in several respects. Firstly, it is not as comprehensive; instead, the premise here is that by concentrating on the three key periods, supplemented by a broader account of the relevant material from the periods in between, we can arrive at a faithful portrait of Indian mathematics as it became deeper and broader with time. Within this general framework, certain themes relating to the posing and solving of important problems and/or departures along new directions can be identified. These milestones, together with an indication of how earlier work fed into them, are shown in the schematic diagram alongside. Consistent with this thematic partition, the approach followed in the book is largely chronological. But there are places where I have had to deviate from strict chronological progression, partly in order to accommodate proofs which are often found in later commentaries, even when we have reason to believe that they were known earlier. More generally, it helps to highlight the continuing relevance of certain themes across very long time spans, for example the influence of the principles of decimal enumeration on several aspects of the work of the Nila school (see the flow diagram). Also, occasionally, I have found it more illuminating to look at the initial steps in the study of a particular topic from the vantage point of how it developed later. An example, out of many, of the advantage of inverting the chronological order is the light Vedic geometry throws on its possible Indus Valley antecedents.

It is not possible in a book of reasonable size to treat every advance in all these areas with equal thoroughness. Choices have had to be made – there was a lot of mathematics in ancient India, and many mathematical texts. As a rule, I have tried to pick for detailed treatment those topics which are central

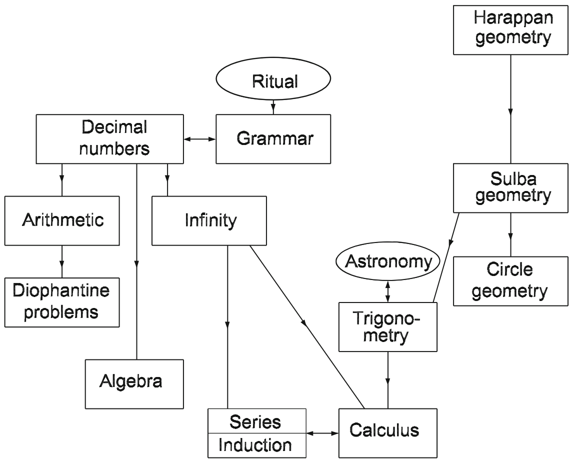

Figure 1: The main themes. Time runs from top to bottom, though not to scale. Oval compartments are for topics not discussed in detail. 'Infinity' is shorthand for the unboundedness of numbers and 'circle geometry' stands for the geometry of cyclic figures, especially quadrilaterals. Combinatorics is missing from the diagram; it derived from prosody and had little direct impact on later mathematics. The attentive reader will probably add a few extra arrows to the diagram.

to the overall development depicted in the flow diagram, the 'big themes': decimal enumeration in the early Vedic texts, Aryabhatan trigonometry, and the invention of calculus by Madhava, to mention only the most prominent. I have also paid special attention to recent work either addressing questions which had not been studied earlier or clarifying issues left ambiguous or unresolved in previous work. Typical examples would be the use of the geometry of intersecting circles in making squares and generating space-filling periodic patterns (Indus Valley), the logic behind Brahmagupta's theorems on cyclic quadrilaterals and the first glimpses of an abstract algebraic mode of thinking in the introduction of general polynomials and rational functions in one variable in the Nila work. Given the interdependence of these themes, it goes without saying that there will be a fair degree of overlap, much to-ing and fro-ing, in their treatment. For the same reason – or due perhaps to the influence of the admirable pedagogic techniques of *Yuktibhāṣā*? – there will also be a certain amount of repetition of some of the key points and themes.

Secondly, as far as the historical aspects are concerned, the approach adopted here is somewhat different from the usual. Due to the efforts of many scholars, we know a little more than we used to, not so long ago, about the lives and work of quite a few of the mathematicians belonging to historical times – basic biographical data, approximate dates and places of composition of key texts and so on – though much remains to be discovered in future work.

Since these facts will (with a few very important exceptions) only get a passing mention, it may be useful to the non-specialist reader to see summarised in one place what is known of the chronology of the major landmarks and their authors, together with the key texts. As far as the early phase is concerned, that is easily done: we have only to remember that the *Ṛgveda* was compiled in the form in which we know it around the 12th century BCE from material composed earlier by anonymous poets (even if names can sometimes be attached to particular poems), the earliest *Śulbasūtra*s (of Baudhayana and Apastamba) in the 8th century BCE and the *Taittirīyasaṃhitā* roughly halfway between the two. Further down the line, it seems safe to place Panini in the 5th century BCE and Pingala a century or two afterwards. All these dates, based mostly on indirect linguistic evidence, reflect the consensus of modern scholarship.

From Aryabhata onwards, we are on much firmer ground. The following list, confined to those who had the greatest impact, either on their followers or on the judgement of modern historians, is only meant as a first orientation. (I have omitted Bakhshali since we know neither the date nor the authorship).

Aryabhata (born 476 CE): His only known work, *Āryabhaṭīya*, is internally dated 499 CE.

Bhaskara I (late 6th - early 7th century): The first prophet of Aryabhata, author of three expository works, among them *Āryabhaṭīyabhāṣya*. All of them had great influence on the spread of the Aryabhatan doctrine.

Brahmagupta (first half of the 7th century): Author of *Brāhmasphuṭasiddhānta* (628 CE), strongly anti-Aryabhata on details of doctrine, and of *Khaṇḍakhādyaka* (about 650 CE), in which he reconciles his astronomical model with that of "master" (*ācārya*) Aryabhata.

Śaṅkaranārāyaṇa and his teacher Govindasvāmi (9th century): The first astronomer-mathematicians authentically established to have worked in Kerala, devoted followers of Aryabhata and Bhaskara I, and precursors of the Nila school. *Laghubhāskarīyavivaraṇa* of the former is dated 869 CE.

Mahāvīra (9th century): His book *Gaṇitasārasaṃgraha* has no astronomy but is strong in arithmetic and mensuration. The only historically identifiable Jaina mathematician and the least Aryabhatan of all.

Bhaskara II or Bhāskarācārya (born 1114): Master of the astronomy and mathematics of his time, a great teacher and consolidator. His *Siddhāntaśiromaṇi* (mid-12th century) incorporates *Līlāvatī* and *Bījagaṇita*.

Narayana (Nārāyaṇa Paṇḍita, 14th century): Accomplished in all aspects of mathematics. His *Gaṇitakaumudī* (1356) has many strikingly original results, especially in combinatorics and cyclic geometry. May have been from Karnataka (or, remotely possibly, from Kerala) which will, perhaps, link him to

Madhava (approximately 1350 - 1420 ?): Astonishingly little is known about the inventor of calculus. One short text attributed to him, on the moon's motion and, maybe, one or two other fragments.

Nilakantha (1444 - 1525-1530): The conscience-keeper of the Nila lineage. Wrote many books of which *Tantrasaṃgraha* had great influence. The masterful *Āryabhaṭīyabhāṣya*, which he himself calls *Mahābhāṣya*, the Great Commentary (like Patanjali's commentary on Panini's *Aṣṭādhyāyī*), written late in his life, is *the* key to Aryabhata's mind.

Jyeshthadeva (second half of 15th - first half of 16th century): Hardly anything known about his life. Disciple of Nilakantha and author of *Yuktibhāṣā*, by far the most important book about the work of Madhava and the Nila school (and, in my view, one of the classics of all time).

The overall picture that emerges is of a coherent mathematcal culture, no matter who contributed to what parts of it, or when, or where. I will return to the theme of the sense of unity all of them together convey at the end (Chapter 14) of this book.

A third issue which is really subsidiary to the main concern of this book but which cannot be avoided is that of possible mathematical contacts between India and other cultures at various times. It cannot be avoided because there has been a trend among historians in the European tradition, from Colebrooke onwards and still quite alive, to hypothesise external influences – Babylonian, Greek, Chinese, etc. – in the genesis of several Indian mathematical developments. This is a difficult issue to be categorical about. Documentary or other unimpeachable evidence in support of (or against) such transmissions does not exist as a rule. Historical plausibility based on cultural contacts at the appropriate time can sometimes be suggested but, often, cannot be built upon. In these circumstances it would seem reasonable that the case for or against transmission is best made by looking at the parallel evolution of a mathematical theme or a related set of themes as a whole, as well as at ideas and techniques that are *not* held in common, rather than at the odd shared factoid, in other words by paying attention to the totality of the internal evidence. On the occasions when this issue does come up (mostly, but not only, in relation to geometry), I have tried to do that. A good case study is the clear and strong Alexandrian, specifically Ptolemaic, influence on Aryabhata's system of astronomy. The internal evidence for a corresponding inflow of purely mathematical ideas is weak and impossible to identify objectively, despite some confident claims made to the contrary. It is perhaps inevitable that, in the reverse direction, there is currently a tendency among a few writers to see European calculus of the 17th century as deriving its initial impetus from the earlier work of the Nila school. Here again, there is so far no evidence in support of the hypothesis. This too is a question I will discuss as it comes up and at the end (Chapter 16.2).

If evidence of a flow of mathematical ideas into India is weak, the case for an organic evolution within India, dating back as far as we can go, is correspondingly strong; the delineation of that continuity is one of the aims of this book. As part of that aim, I have tried to frame Indian mathematics within the general context of Indian history and culture, of the movement of people and ideas, linkages with other disciplines, in short the intellectual climate of the times. The historical interludes that are occasionally sandwiched among the descriptions of the mathematics are meant to serve that purpose. Apart from providing the general context, it is a fact, at least in India, that a broader cultural view often throws a direct light on the substance and style of the mathematics itself, though it is understood that culture and history alone never produced a genius. Like all historical reconstructions, such an endeavour cannot be anything like the last word; new facts and fresh insights will continue to emerge and bring with them the need to reevaluate some of the connections tentatively suggested here.

That having been said, it bears repetition that this book is really and fundamentally about the mathematics. In particular, even astronomy *per se*, which was the original motivating force for the Aryabhatan revolution, will only get passing mentions.

In the manner of presenting the mathematical content, the historian is faced with the eternal dilemma of all history of science: finding the right balance between literal faithfulness to the sources and the wish to see what was done in the distant past from the perspective of what we consider interesting and important today. Absolute fidelity to the original texts risks losing the present day reader, used as he or she is to the habits of thought and discourse of a science, mathematics in the present case, that has grown unrecognisably in depth, generality and sophistication in the last few centuries. It is of course axiomatic that a literal reading of texts must remain the foundation of whatever conclusions one may draw from them and good translations have precisely that function, leaving the reader to make the adjustments necessary to fit these into the mathematical universe that we inhabit now. For European mathematics, this has been relatively easy to manage as the modern style of mathematical writing is in a direct line of descent from the Mediterranean-Greek. There are in fact successful books which rely on that approach. Such total literalness is less likely to succeed with Indian material, which has its own style of communication – as already noted in the section on sources – indeed its own way of thinking of and doing mathematics. Most scholars of the Indian tradition, from Colebrooke down, have understood this, at least implicitly. Several recent studies have in fact taken the route of accompanying a translation of the work concerned, also given alongside in the original language, with detailed notes in the modern style and in modern terminology and notation: Sarma (*Yuktibhāṣā*), Patte (commentaries on *Līlāvatī* and *Bījagaṇita*), Ramasubramanian and Sriram (*Tantrasaṃgraha*), to cite only the most recent.

In a general book such as this, aimed mainly at interested but not specialist readers, this scholarly ideal is difficult to sustain even for selected extracts

and perhaps not desirable. In its place, I have adopted a less demanding approach, limiting direct citations from the texts to translations of a few (very few) essential or insightful passages and freely using symbols, equations, diagrams and terminology that are the currency of modern mathematical communication at all levels, while trying at the same time to remain faithful to the original content; very occasionally and mainly as context, there will even be evocations of more modern concepts. Such an approach also helps to counter the temptation of looking at Indian mathematics with its non-Greek pedigree as something exotic and alien, done by people with strange ways of thinking, to be understood after strenuously mastering strange and difficult languages. The hope is that the reader will see at the end that the mathematics of India is the same mathematics that he or she is used to – and that applies also to Indian readers. It goes without saying that some of the mathematics so conveyed will be elementary, school and college mathematics of present times. But there will also be topics which, inherently and in their treatment, we will consider today to have been sophisticated and modern in spirit. Examples would include Aryabhata's introduction of functional differences, Brahmagupta's method of generating new solutions of the quadratic Diophantine equation from known ones, the quasi-axiomatic treatment in *Yuktibhāṣā* of the set of positive integers, possibly as preparation for inductive proofs, and many others – modernity is not necessarily an increasing function of time. An abstract notation and the mindset that goes with it are an advantage in assigning such advances their proper place in the evolutionary chain of mathematics as a whole. At the same time, and equally usefully, the detachment that results from a degree of abstraction will let us see more clearly the difficulty a verbally expressed mathematical culture faces in extending and generalising knowledge acquired in a particular context.

But translating mathematical verse and prose into an abstract-symbolic language also has its risks. It is all too easy to read more into a piece of text than was actually there, e.g., to see in the equation representing a verbal passage all that that equation conveys to the prepared mind of today. There is also the opposite risk, that of devaluing the originality and power of an idea or method because everything that a modern mathematical sensibility expects to read into its symbolic expression is not to be found in its first formulation. Examples of both kinds of misreading are easy to come by in the literature, as we shall have occasion to note. A conscious effort to keep out the biases of modernity is nevertheless worth making, even if destined to be not wholly successful.

This is not meant to be a scholarly book. It began life as the notes for a course given to university students of science and its primary goal is to inform and instruct those who, like my students, wish to acquire a general picture of the mathematical culture of India that does not sacrifice essential pedagogic details; all mathematical statements made are proved and the proofs are those in the texts, occasionally streamlined, with notice given, to satisfy the modern norms of mathematical writing. My aim of keeping it reasonably self-contained is also reflected in the way I have selected and organised the bibliographic

material. It is by no means exhaustive. There are lists at the end of the book of the indispensable original texts (including translations into English when available) and of secondary sources that I have found most useful. References that relate to specific points are relegated to (infrequent) running footnotes.

0.4 Sanskrit and its Syllabary

Though, in the interests of a general readership, I have avoided quotations in the original Sanskrit, it is difficult to keep out Sanskrit words and phrases altogether, and for more than one reason. There is first the need to convince that translations are verifiably accurate. When more than one way of rendering a particular term suggests itself, or has been suggested in the literature, I have felt it useful to accompany my chosen reading by the original expression. That is a small price to pay for keeping away from inelegancies such as the literal 'seed calculation' (*bījagaṇita*) for algebra or 'debt amount' (*ṛṇasaṃkhyā*) for a negative number. Students of Sanskrit know that it is a precise language and that possible ambiguities can often be resolved by a rigorous reading of the exact terminology employed. Indeed, problems of great mathematical and historical interest which revolve around the phraseology used, and can be resolved by attending to the rules leading to that phraseology, are not unknown.

At the same time, there are also places where different words are used to connote the same concept or, conversely and more problematically, one word may stand for distinct though related concepts. Such deviations from a one-one correspondence between mathematical objects and operations on the one hand and terminology on the other is to be expected in writing that is entirely in narrative form in a natural language. But it can be confusing. The problem afflicts Aryabhata – who's *sūtra* format has little room for precisely constructed names and their explanations – as much as it does the work of the Nila school where terminological inventiveness has a hard time keeping up with the proliferation of innovative ideas and methods.

Related to these is the recourse to certain formulaic word combinations which are designed to convey much more than their literal meaning. The sense to be attributed to them is decided either by convention or the mathematical context in which they occur. Perhaps the most common is *trairāśikam* ('of three numbers') generally translated as the 'rule of three', the rule which determines any number out of four, which are in equal proportion pairwise, from the other three. Even more rich in content is the Nila school's *jīveparasparanyāyam*, 'the principle of adjacent chords', which stands for the addition formula for the sine function *and* Madhava's derivation of it. In neither of these cases is the full meaning of the formula in doubt but that is not always true.

These are all good reasons for the distracting presence of the occasional Sanskrit phrase. Mainly they serve as guideposts to help us stay on the straight and narrow path of fidelity to the texts, especially in a book such as this that is twice removed from the original material, translation into English followed

by a rephrasing in modern mathematical terms. There is however one area into which it is virtually impossible to venture without getting involved in the linguistic and grammatical subtleties of Sanskrit – where the mathematics is in the language itself – namely the development of the decimal place-value system of numeration. The precise quantification or 'measurement' of a general number (the cardinality of an arbitrary finite set) by means of 10 as base in early Vedic times came about orally, not through symbols for the numerals up to 9 and their relative positions in a written representation, but through a choice of names for these numerals as well as for the 'places', the powers of 10. To get the decimally determined name for any number, the primary number names had to be combined following the grammatical rules of Vedic (and, by and large, later) Sanskrit, the rules of nominal composition. What is a simple and easily grasped matter of ordering in a written positional notation is thus subsumed in the elaborate rules governing the composition of words and syllables. I have tried to keep grammatical excursions to the essential minimum, but a feeling for these rules is of great help in recognising the mathematical role they played in Indian decimal enumeration. Indeed, the early symbiosis between numbers and words remained a potent influence on both grammar and arithmetic for a long time.

The Sanskrit words and phrases that are present in this book are written in Roman script, making use of diacritical marks according to standard (IAST) practice. Guides to the way the marks determine pronunciation can be found easily (on many Internet sites for example). But since the syllabary they encode follows certain principles of organisation which have had at least two mathematical manifestations – Pingala (Piṅgala)'s combinatorial approach to prosody and a particular syllabic transcription of numbers, the so-called *kaṭapayādi* enumeration (we will come back to them in due course) – I use this occasion to describe its organisation very briefly. In any case, it is very much a part of the history of knowledge in our first phase; there are in fact scholars who have expressed the view that the science of language in all its aspects is the first science of India.

The currently accepted picture of the history of the Vedic people is that sections of them moved east and south over several centuries, eventually inhabiting most of the Gangetic plain. By about the 5th century BCE, they had a strong presence as far east as present day Bihar and it is to this time and place that the fixing of the final form of the Sanskrit syllabary is credited. It has remained virtually unchanged to the present, and has in fact become, with very minor changes, the syllabary of the majority of the main modern Indian languages (though there are marked differences between them in their written form).

We have, first, the vowels:

a ā i ī u ū e ai o au

In the pronunciation, the unbarred *a*, *i* and *u* have the same duration, called 'short', while their barred variants, called 'long', have twice that duration. The

diphthongs *ai* and *au* are long, consistent with how we articulate them, but the surprise is that *e* and *o* are also long in Sanskrit (no bar on them since the short variants never occur). I add for future reference that in Malayalam, the language of *Yuktibhāṣā*, *e* and *o* do occur in both short and long versions like the other vowels, the long being distinguished, in that context, by an overbar.

To these have to be added the 'semivowels' *ṛ* and *ḷ* which come after *ū*, as well as the 'seminasal' *anusvāra* (*ṃ*) and the 'semi-aspirate' *visarga* (*ḥ*) at the end of the list.

Apart from being syllables in their own right, the vowels make possible the vocalisation of the consonants which also therefore come in two durations. The main body of the consonants (conventionally vocalised with the short *a*, any other vowel will do as well) is arranged as a grid, the 'consonant square' of five groups of five syllables each:

ka	*kha*	*ga*	*gha*	*ṅa*
ca	*cha*	*ja*	*jha*	*ña*
ṭa	*ṭha*	*ḍa*	*ḍha*	*ṇa*
ta	*tha*	*da*	*dha*	*na*
pa	*pha*	*ba*	*bha*	*ma*

The syllables are to be read in the normal sequence, first horizontally and then vertically. Apart from these twentyfive, there are nine sundry consonants which do not fit into the grid and so are appended at the end:

ya *ra* *la* *va* *śa* *ṣa* *sa* *ha* (plus *ḷa* in Malayalam).

The point about the grid is that as we proceed down from the top row, the vocalising point for the production of the corresponding sound moves progressively from the farthest part of the palate towards the front, to the teeth and then the lips. The syllables with the underdot (the middle row as well as *ḷa*) are all retroflex, for example. And as we proceed from the extreme left column to the right, the primary sound (*ka*, *ca*, etc.) acquires a modulation depending on the simultaneous use of the breath or on minor changes in the sound-producing action. Thus the second and the fourth column are the aspirates and the fifth the nasals.

A compound consonant (generally of two 'pure' consonants, occasionally of three, as for example in *matsya, śāstra*) is one syllable, pronounced with only the final consonant (*ya, ra* in the examples) having a vowel value, the preceding ones being employed without vowel value or duration. This is not specific to Sanskrit (though there are Indian languages which do not follow this practice). It is also specified that the syllable preceding a compound consonant is to be pronounced long even when it is, in isolation, short. This only codifies something we know from actual speech: the duration of each syllable in the pronunciations of *matsya* and *māsa* for instance is the same even though one word has the short *ma* and the other the long *mā*.

In the interest of readability, I have avoided the use of diacritical marks on the names of people and places, but have indicated the correct pronunciation in brackets on their first occurrence when it was felt to be useful. The names of individual texts are generally given in full diacritical glory since often their meanings do matter.

Part I

Beginnings

1

Background: Culture and Language

1.1 The Indus Valley Civilisation

The earliest civilisation of India that can be reconstructed in a reasonably coherent manner is the Indus Valley or Harappan civilisation. We have a pretty accurate idea of its geographical extent and of its duration in time, almost entirely from archaeological evidence. In its pomp, it extended over almost all of modern Pakistan, i.e., the basin of the Indus river system, as well as southeastern Afghanistan, the northern half of the Indian state of Gujarat and the area between the tributaries of the Indus and the river Yamuna to the east. Of the ancient river valley civilisations, it covered by far the largest area. It used to be thought that its beginnings date from around 3,000 BCE but modern archaeology has traced its roots to agricultural settlements of the late stone age going back very much earlier, to the 7th millennium BCE, in an area around Mehrgarh in Pakistan, south and east of the Bolan pass. Archaeologists now characterise the period from ca. 2600 BCE to ca. 1800 BCE as its mature phase, sandwiched between a proto-Indus early phase and a late or declining phase. Though clear cultural continuities can be seen across the transitions between these loosely defined periods (it must be kept in mind that, despite much recent progress in scientific dating methods, the dates cited are, generally, indicative), it is primarily the mature phase that is of interest here. This is the time when urban settlements on a large scale, planned with a high degree of sophistication, came into being. The ruins of these cities and the rich and varied collection of artefacts found in them are the main source of what we know about the Indus Valley culture. The focus of the extremely brief and qualitative summary in this section of its defining characteristics will therefore be confined to this phase.

Given the immense geographical reach of the Indus civilisation, the most remarkable of its many remarkable features is the uniformity of the cultural

P. P. Divakaran, *The Mathematics of India*, Sources and Studies in the History of Mathematics and Physical Sciences, https://doi.org/10.1007/978-981-13-1774-3_1

profile that archaeologists have drawn from its remains. Urban habitations are found in all parts of the territory, at least five of them having an area and a population that qualify them as proper cities, from Harappa and Rakhigarhi (north of Delhi, yet to be thoroughly excavated) to Mohenjo Daro on the lower Indus and Dholavira and Lothal in Gujarat. The inner city plans were similar, generally rectilinear, with straight roads intersecting at right angles and aligned in the cardinal directions. A great deal of the construction was in fired bricks of standard dimensions and of fine geometrical and material quality. Architecture was also by and large planar with rectangular floor plans and elevations, again with a high degree of accuracy. Weights and linear measures follow the same progression wherever they have been found. Most surprisingly, the samples of writing on seals – perhaps *the* defining feature of the culture, as inscribed clay tablets are of Mesopotamia – and other media do not change from region to region, even if we do not know whether they always expressed the same language.

The other striking feature is the absence in the entire domain of any signs of great authority being exercised – nothing that evokes a powerful royalty, no ostentatious palaces for example – or of violent activity on a large scale, such as wars and uprisings. The general impression given is that of a prosperous and peaceful society that valued order in all the things they did, going about its daily business of agriculture, manufacture and commerce. Wealth-generating trade in particular, both domestic and with other contemporary societies, was definitely an important part of economic life: Indus seals have been unearthed in Mesopotamia for instance and there are mentions in Mesopotamian inscriptions of a country called Meluhha, very likely the coastal Indus region, from where exotic commodities were imported.

It seems reasonable to suppose, as many historians do, that the sustenance of a materially advanced society over long periods must be based on an understanding and exploitation of what is now called science and technology. In ancient times for example, trade in any form other than face-to-face barter, and especially when the counterparty is thousands of kilometers away, could not have been carried out without accurate methods of quantification of the goods being exchanged, both in numbers and in volumes and weights. Thus, counting and measurement, and doing basic arithmetical manipulations with the resulting numbers, are skills the Indus people could not have done without. Indeed, the few thousand examples of written symbols discovered so far include signs – vertical lines grouped together – which almost certainly stand for numerals. And among the most beautiful of the Indus artefacts are the weights, perfectly cubic and made of polished soft stones like chert, doubling in sequence at the lower end of the scale and following a more complicated order involving binary and decimal multiples as they get heavier.

Trade also brings in the question of the scientific knowledge involved in long distance travel. There is now a consensus that the Mesopotamian trade was partly sea-borne, from the mouth of the Indus and the Gulf of Cambay (Khambat, in Gujarat) to the mouth of the Euphrates-Tigris (there was also a

link with northeastern Oman which was on the way); a large basin in Lothal is thought to have been a dockyard and Indus artefacts have been found near Bahrain. If that was the case, the mariners would have had to have an adequate knowledge of celestial navigation. Quite independently of the navigational demands of long voyages, settled agriculture, for which evidence exists from the earliest strata of Mehrgarh (ca. 7000 BCE), would have initiated the monitoring of seasonal changes and their relationship with celestial phenomena.

The other area of mathematical interest we can hope to learn something about is geometry. Here the inferences that can be drawn are somewhat stronger, primarily based on a study of the architecture (at several scales) and of some beautiful decorative patterns, incised or painted on seals and sealings or on potsherds. How were the cardinal directions determined? What methods were used by surveyors to make straight roads intersect at right angles and by builders to design rectangular floor plans and elevations? What were the geometric principles in play in the generation of ornamental patterns as regular lattices from a 'unit cell' and how was the unit cell itself drawn?

What makes the answers to such questions inferential rather than a matter of certainty is the fact that the Indus writing has so far remained undeciphered. The script is almost surely logographic, a system in which each symbol stands for a word, with an admixture of purely syllabic signs (though no one has been able to separate the two types). There are between 400 and 500 such symbols, the exact number depending on how we decide which of them are independent and which are (grammatically determined?) modifications of one another. The samples of writing that have survived (it goes without saying) are on durable materials like the stone of which the seals were made or the clay of potsherds. The carving is of an exceptionally high standard and often accompanies equally finely rendered depictions of animals, vegetation, ritual objects and humans both naturalistic and 'god-like'. There is a high degree of symmetry in the individual characters, about 70% of them exhibiting bilateral symmetry, generally about a vertical line.

The writing is invariably horizontal. We also know that it is from right to left, a fact first surmised from the mechanics of the incisions and recently confirmed by statistical analyses. Sentences are relatively short as is natural in a logographic script - there is more content conveyed in one logographic symbol than in a symbol of an alphabetic (European languages) or a syllabic (Sanskrit) script. (English, for instance, has only three one-letter words, at least until texting became widespread). To compensate, there has to be a much larger number of symbols, one for each word, than in the more abstract syllabic or alphabetic representation of spoken sounds; a few hundred are generally thought to be adequate for basic communication.

Issues such as these are important because a structured language would appear to be an essential prerequisite for the development of scientific and mathematical thought. To add to the relevance of the question, there has recently been some work doubting the credentials of the Indus Valley inscriptions as embodying a natural language, on the basis of a rather arbitrary set of criteria

for a set of ordered symbols ('writing') to encode a language as well as a superficial look at some statistical correlations. As sometimes happens in science, the resulting uproar among Indus scholars led to much deeper analyses of syntactic correlations, something that could have been done decades back. It turns out that the scripts of natural languages have certan statistical structural properties which are approximately universal, independent of the language and the script in which it is written, and that the Indus script, encoding an unknown language, shares these characteristics. It would seem that the Indus people really had a language.

Attempts to decipher that language have been going on now for close to a century, with very little to show for it. The least unlikely of the suggestions made so far – the double negative is a measure of the intractability of the problem – is that it is the ancestor of a putative language, early Dravidian or proto-Dravidian, from which, in turn, Old Tamil descended. Such a possibility is of some interest in regard to the evolution of counting, as we will see in Chapter 4. What is somewhat puzzling is that, apart from a small pocket in Baluchistan speaking a language (Brahui) of the Dravidian family and apart from a number of loan words from Old Tamil in early Sanskrit, there is no evidence for a living and vigorous presence of Dravidian languages in the core Indus territory in historical times. The separation may have been gradual, but it also appears to have been irreversible.

Very recent discoveries of fragmentary inscriptions in Indus-like characters in Tamil Nadu and Kerala provide some tenuous (as of now) backing for the idea that there was a written proto-Dravidian language with links to the Indus culture. For a language to migrate from the northwest of India to its extreme south so as to change the linguistic landscape so completely, both in its place of origin and the place it ended up in, there must have been substantial movements of the people speaking that language. That the bigger Indus cities like Mohenjo Daro and Harappa began to lose their populations towards the conclusion of the mature phase had been surmised already by the pioneers of Harappan archaeology. More recent work tends to show that this happened over the entire Indus domain, relatively rapidly; by about 1,500 BCE, the urban centres became depopulated, compensated by an increase in the number of much smaller settlements, in contiguous areas, far less urban and orderly and with very few of the hallmarks of an Indus city in its glory such as sophisticated civil works or the profusion of fired bricks. Even that seems to have come to an end a century or two later, giving way to an apparently new culture that then spread slowly eastwards. It is perfectly possible that some of the hardier people took the route south, either by sea along the coast – they had a strong maritime tradition after all – or overland.

What reasons compelled the Indus Valley people to uproot themselves from their settled and prosperous home and set out on a long journey in search of an uncertain new life? The consensus of opinion is that the direct cause of the emptying of resource-consuming cities and the migration away from the Indus heartland as a whole was environmental disturbance of some sort or the other:

climatic changes and the accompanying disruption in the pattern of rainfall, the drying up of the Ghaggar-Hakra river (the mythical Sarasvati) which arose to the east of the Indus river system and used to flow along a parallel course, etc.[1] Whatever the reason – and civilisations do degenerate by themselves – the disappearance of the Indus civilisation left a vacuum which seems to have been filled almost immediately by the Aryan-Vedic people. And the first geographical area settled by this newly prominent culture was none other, more or less, than the northern part of the Indus homeland.

The transition from the Harappan to the Vedic marks a period of almost total discontinuity in Indian cultural history. Cities and towns disappeared from the landscape and the technology of fired bricks seems to have got lost except, intriguingly, in the making of the ritual fire altars of the Vedic people. These were small in dimension (as attested by the descriptions in the *Śulbasūtras*) when compared to the grand Indus Valley buildings, and few in number. We have to wait a thousand years and more to see the revival of these ancient technologies, on a greatly diminished scale, far to the east along the Gangetic basin and in Gandhara. Much more drastic is the change in the nature of literacy. Logographic writing vanished from India, never to return, not even in the Dravidian south; indeed, it is widely thought that the early Vedic people did not have any writing at all. When writing does make a reappearance, in the rock inscriptions of Ashoka (Aśoka) in the 3rd century BCE, it is syllabic – there are a few scattered inscriptions from deep southern India in a script having an affinity with some Indus characters – as it has been in all Indian languages ever since. And, from that time, the direction of writing in all of India has been from left to right. (There is a minor exception to this, Kharoshti; see below, section 3). How can we account for these dramatic shifts in the most fundamental of the traits that define the cultural identity of a people? There is no convincing answer. Every little scrap of reliable information that can help bridge this gap will have immense value.[2]

1.2 The Vedic Period

As far as material remains go, about all that the early Vedic age has bequeathed to us is a distinctive style of pottery, less refined than its Harappan predecessor, supplemented a little later by some iron implements. Aside from serving as markers of the transition to the Vedic period (as well as between the copper-bronze and the iron ages) and of geographical territory, they do not provide much insight into the cultural identity of the people, certainly not as regards mathematical knowledge. Virtually everything that is known about the early

---

[1] A good way to keep track of continuing work in the field, in both senses, is to follow the website www.harappa.com.

[2] A recent excavation in Kerala has exposed structures made of fired bricks some of which have the same dimensions as the Indus Valley 'binary' bricks, see below. If nothing else, they bring some concrete support to the hypothesis of the southern migration.

Vedic people comes from the texts that they created in such abundance; and there is every reason to think, as we shall see, that these texts have come down to us essentially unscathed by the passage of time.

From the perspective of intellectual developments, a natural time frame for the Vedic age is provided by the collation of the poems or hymns of the *Ṛgveda* into ten circles or chapters (*maṇḍala*, 'Books') as the initial point and, roughly, the middle of the first millennium BCE as its terminal point. Mostly on the basis of the study of its language, it is widely held that the gathering together of *Ṛgveda* took place in ca. 1,200 - 1,100 BCE, five or six centuries after the arrival of a new ('Aryan') culture in northwest India. The individual poems of the so-called family Books (Books II to VIII) were composed during this time. These are dominantly poetic laudations of the main Vedic gods and invocations of their power to grant boons, with an element of mystic vision running through many of the hymns. The visionary quality is much more pronounced in the late Books I and X as also, appropriately, in Book IX concerned mostly with Soma, the god of the Vedic drink of the same name, supposedly intoxicating or hallucinogenic. The late Books are not only distinctive in their thematic content and poetic quality – the best-known Vedic quotations come from them – but also, according to experts, linguistically. (Such distinctions are of some interest in the context of dating the genesis of decimal place-value enumeration).

The language of the *Ṛgveda* is Sanskrit in an early form. Some scholars prefer to think of it as a distinct language called Vedic, but its evolution into the later Vedic and subsequently into the early classical form of Sanskrit is direct and continuous – most of Panini's grammatical principles apply to the *Ṛgveda*, for example. As for its own origins, there is no serious reason to question the view of William Jones, already at the time of his rediscovery of the *Ṛgveda*, that it is a descendant, along with a multitude other languages spoken in parts of Asia and Europe, of a hypothetical language, Old Indo-European; work over the past two centuries and a half has only strengthened the conclusion. But philologists have also identified, more recently, some non-Indo-European affiliations in the form of loan words from Indian language reservoirs such as the Munda of the indigenous populations of north India and, very surprisingly, Old Tamil. The *Ṛgveda* and the unexpectedly bright light it throws on the numerical culture of the earliest Vedic people will occupy us at some length later.

Soon after came the *Yajurveda* concerned, mostly, with ritual practice. There are two versions of it of which the earlier, the *Taittirīyasaṃhitā*, will be of interest to us. Two more texts, the *Sāmaveda* and the *Atharvaveda*, complete the quartet of the so-called *saṃhitā* texts, the original Vedic corpus. Apart from the content, the structured verse format in which they are wholly (e.g. the *Ṛgveda*) or partly composed, bound by rules regarding syllabic sequences and lengths, is itself a window to the early Vedic understanding of the combinatorics involved. Each of these four texts generated in time a line of explanations and interpretations of the originals, the Brahmanas (*Brāhmaṇa*) and the Aranyakas (*Āraṇyaka*), some of which also contain nuggets of mathematical knowledge, buried often in a lot of obscure elaboration.

In the next category of texts, the Upanishads (*Upaniṣad*), we can begin to see a gradual change in intellectual preoccupations, a return to the philosophical concerns of the later books of the *Ṛgveda* if not always to their poetic vision. The word *upaniṣad* means 'sitting together', as when a guru initiates a disciple into some esoteric knowledge; the confidentiality evoked hints perhaps at an attitude of scepticism and rebellion, of going against received dogma. In any case, the emphasis has begun to move away from faithful mechanical practice of ritual to matters of the mind: earlier positions are questioned, not only as regards ritual doctrine but also their metaphysical basis. Controversy and debate became, if not immediately respectable, at least part of the discourse. It is not without reason that some of the main Upanishads came to be clubbed together as *Vedānta*, the end of the Vedas. It is also not a matter of chance that out of this churning arose the new faiths of Jainism and Buddhism with their stress on a rational understanding of this world in all its aspects.

These facts alone justify the choice of mid-first millennium BCE as marking the approximate closure of the Vedic period. But several other trends also converged around this time to produce a cultural transformation, a shift towards intellectual enquiry, whose causes and consequences were not limited to spiritual matters; even the two new religions had strong metaphysical and even cosmogonic underpinnings. And, in the secular sphere, there were fresh analytically oriented approaches to old and new problems. Philosophical systems were identified and classified, rules of logic and reasoning, even the regulations governing debate and disputation, were subjected to scrutiny, not least in the teachings of the Buddha and Mahavira (the founder of Jainism) themselves. Language and the scientific rules that govern its use took centre-stage: the consonant square got its final form; the way compound words are to be broken up (word-by-word recitation, *padapāṭha*), both in order to exhibit the grammatical rules of the phonetic transformations involved and to freeze the correct pronunciation of texts, was systematised. Above all, the "Eight Chapters" (*Aṣṭādhyāyī*) of Panini formalised the grammar of the Sanskrit of his place and time (and, almost, for all time to come), followed a century or two later by Patanjali's "Great Commentary" (*Mahābhāṣya*) on it. All this led on, within perhaps three centuries after Panini, to the singularly original work of Pingala on prosody in which he constructed a mathematical classification of all distinct metrical schemes as sequences of long and short syllables of a fixed syllabic count.

This focus on making precise the study of Sanskrit was preceded by a less analytic, or more amorphous, concern with language in its various aspects. Some time after the early Vedic era, beginning about 800 BCE say, there came to be composed a set of manuals called *Vedāṅga*, the 'limbs of the Veda' (of which we know six) dealing with *śāstra*, the 'sciences'. It is quite clear that their purpose was to anchor the understanding of Vedic practices to the knowledge base of the priestly class. What is remarkable is that four out of the six sciences are concerned with the study of language: phonetics or phonology, etymology, grammar and prosody. Equally interesting is the first limb which deals with

the science of ritual and of which one part forms the *Śulba* texts in their early recensions, the workbooks on the architecture of ritual structures in brick and the geometry it is based on. Astronomy (*jyotiṣa*) brings up the rear and it is generally about time determination and the calendar, presumably as a guide to the choice of the appropriate time for ritual-religious observances. Not only was geometry thought of in purely architectural terms, having little to do with astronomy – we have to wait till Aryabhata and the Siddhanta period for them to be brought together to such good effect – its role was as only one of the several agents that contributed to making ritual activity free from error.

Historically, our knowledge of the early Vedic period is very meagre. About all that is known with reasonable confidence is that the earliest Vedic people settled in the northwest and then moved east, to Panjab at the time of the compilation of the *Ṛgveda*. With time, the culture spread, first through the Kuru country (north and east of Delhi) and later along the course of Yamuna-Ganga until, by about 500 BCE, the entire river plain and the areas north of it, at least down to modern Bihar, became culturally Vedic – indeed many of the philosophical and linguistic innovations mentioned above arose at the eastern edge. Politically, almost all of north India was in the process of becoming a land of many essentially independent entities, ruled over by chieftains or kings and collectives of representatives (*janapada*) of clans, but politics does not seem to have inhibited scholarly exchanges in any way. Especially noteworthy is the fact that the extreme northwest, including parts of southern Afghanistan, remained a vital part of this vast, culturally homogeneous territory, Aryavarta (Āryāvarta) as it was called. Roughly at the time that Shakalya (Śākalya) was busy canonising the spoken and chanted Veda (through the *padapāṭha*) at the eastern extremity of the territory, Panini was occupied, two thousand and more kilometers to the west, in describing the grammar of Sanskrit "as spoken in his time in Aryavarta" (according to Patanjali). Nor did the fact that, in the 6th century BCE, India had experienced its first recorded armed invasion from outside – that of Cyrus of Persia – seem to have had much of a negative impact outside the local power structure. Cyrus' empire extended to the Indus river over a considerable time until, in fact, the next invasion of the same region by Alexander of Macedonia in the 4th century BCE. Taxila, right in the middle of the occupied territory, became a centre of learning during this time and grew into a university (and trading) town that remained vibrantly alive for the next millennium.

It is useful to remind ourselves once again that the new religions of Buddhism and Jainism, both of them giving emphasis to thinking things through as opposed to blindly following long-established dogma, are emblematic of the great transformation that began to unfold around 600 - 500 BCE. Born out of the intellectual ferment symbolised by the Upanishads, they quickly attracted a broad adherence, intially in the eastern region of their birth, followed soon after by a dissemination into other parts of India (and, in the case of Buddhism, also outside). The really dramatic events in this awakening of the spirit of rational enquiry and analysis all took place in the three or four centuries that

followed. Interestingly, this relatively brief period also saw the rejuvenation and readaptation of the old Vedism in a form that is recognisable as the Hinduism of today. The great Hindu epics Mahabharata (*Mahābhārata*) and Ramayana (*Rāmāyaṇa*) were composed in their final form towards the end of this period, the former probably first.

As a final note, it is time to face the question that has so far remained unacknowledged: who really were the Vedic people? Taking into account two undisputed facts, namely the decline of the sophisticated and urban-oriented Indus Valley civilisation in the first half of the second millennium BCE and the temporal and geographical overlap of late Indus and very early Vedic cultures, there can only be two answers: either the Indus people, after having deserted their magnificent cities, reinvented themselves as pastoral-agricultural communities, or the Vedic people moved in from elsewhere and filled the void left behind by a collapsing old civilisation. Which of these alternative hypotheses squares with the scanty and circumstantial evidence available has been the subject of a debate that has been kept alive for some time now. We have seen that there is a sharp discontinuity across the two cultures both materially and, even given that the Indus language is unknown and the script phonetically unread, linguistically. The break in the nature of the artefacts may be capable one day of a plausible explanation without invoking wholesale population shifts. But a language is hard-wired into the people who own it, it is part of their collective mind and it goes where they go. We cannot envisage the relatively quick emergence of a new fully formed language in northern India, so completely different from that of the Indus and other local cultures except, again, as a result of the movement of people. As for where the speakers of Vedic Sanskrit originally belonged, any proposed answer will have to accommodate its many deep affinities with other languages descended from Old Indo-European and the resulting chronological and historical constraints. The old Aryan invasion theories having been discredited, a new model has to be constructed. Where did they come from? Did they arrive all at once or in small waves over a period? In what numbers, and what was the proportion of males to females (so that we may get an idea of the extent of interbreeding that may have taken place after their arrival in India)? Fortunately, these are questions which can be addressed by a careful application of the new discipline of the comparative genetics of populations. The case for a vigorous pursuit of such studies is a strong one.

1.3 The Oral Tradition

The earliest surviving samples of readable writing in India are the edicts of the Mauryan emperor Ashoka (Aśoka) who reigned from 269 to 231 BCE. As of today, more than 30 Ashokan inscriptions are known from all parts of India and the envelope of their find places can well serve as an approximate delineation of the frontiers of the Mauryan empire at its peak. They were inscribed on natural rock formations, on the walls of caves and on pillars cut and erected for

the purpose. One, in Kandahar (Afghanistan), is bilingual, the two languages being Aramaic (the language of Persia at the time) and Greek (Alexander of Macedonia had been in Afghanistan just a century earlier). Apart from that, the inscriptions on the western boundaries of the empire are in Kharoshti, a script that had a short-lived (as scripts go) currency in a region extending from Afghanistan to Gandhara. It was partly inspired by the semitic scripts of the western parts of the Persian empire and was, consequently, written from right to left – the only such to have lodged in India after the disappearance of the Harappan. All other Ashokan edicts are in a regular and beautifully formed Brahmi script, syllabic and written from left to right. They are well composed and the longer ones are quite extensive. Considering that these edicts cover pretty much all of India and that they are in the same script and in popular local variants of Prakrit, not classical Sanskrit, we have to conclude that i) the ability to read was not confined to the erudite classes versed in Sanskrit and ii) Brahmi in a fairly evolved form was read and understood in all parts of India except perhaps in the northwest – there are in fact scattered bits of evidence from the southernmost part of India, outside Ashoka's realm and not much after his reign, for the prevalence of a related script there.

New writing systems are not generally invented on the spur of the moment by some inspired individual and we certainly have no reason to think that Brahmi was put together by a committee as happened with the Korean language in the 15th century CE. It is entirely reasonable to suppose that its evolution from the first tentative scratchings to the mature form which we see all over India in the Ashokan edicts took time, perhaps several centuries of experimentation and diffusion. Whatever was written down during this period of incubation was, by implication, on material that has not survived the passage of time; as Sarasvati Amma says ([SA]), "India's sands were never so kind to her records as Babylonia's sands have been to her clay tablets". Ashoka was clearly addressing posterity, not just his subjects of the day, when he caused the ethical principles on which his statecraft was based to be proclaimed in imperishable rock inscriptions. It is very possible, as many historians believe, that he was inspired to do so by the example of Darius I of Persia, another 'Universal Monarch' who wished the glory of the Achaemenid empire to be recorded in the timeless medium of stone.

But could Brahmi have begun life as far back as, say, 1,000 BCE, 800 years before Ashoka? The accepted answer is 'no'. There are, however, reasons for favouring a more cautious and nuanced response. First, it is somewhat implausible that the highly (orally) literate[3] Vedic people were totally ignorant of the art of writing, living as they did in the wake of a civilisation which did have writing, the Harappan, that was then in the last stages of its collapse. Then there exist sporadic references, generally derogatory, to the skill of the

[3]I will often use the word 'literate' to qualify those who created verbal compositions which were not necessarily written down; otherwise 'Vedic literature' will have to be considered an oxymoron.

script in some middle-Vedic texts[4] as well as an isolated mention of a script and scribes in Panini. Could that script have been some sort of Proto-Brahmi? Still later, possibly in the last centuries BCE, one of the mythical biographies of the Buddha relates a story (which we will have occasion to revisit in connection with very large numbers) in which the prince Siddhartha is depicted as an expert in sixtyfour different scripts. Can it be that writing was always around but was regarded as alien and impure and so disdained by the Vedic seers as a means of communicating their scriptural corpus? To such questions, we do not even have a basis for a good guess at the answers. What seems reasonably certain is that the linguists who organised the sounds of Sanskrit into a systematic syllabary on the basis of phonetic order could not have had anything to do with the choice of Brahmi as their symbolic representation. There is no element of continuity in the shapes of the symbols as we go through the consonant square by row or by column; visually they are not grouped together as they are aurally and the whole thing is as random and chaotic as the alphabet of European languages. Perhaps it was an import grafted on to a phonetically well-organised syllabary, cut into stone to guarantee indestructibility.

Setting aside futile speculation, we can instead focus on the many indications, indirect but nevertheless strong, that point to early Vedic literature being composed orally, then memorised and, finally, passed on orally. The most obvious of them is the stress on phonetic correctness in the use of language, the insistence that the expression of a language is defined primarily by the way it is pronounced, already evident in the scientific organisation of the syllables as determined by the region of production of spoken sounds (the grid of consonants) and in the precise differentiation of the *mātra*, the duration of short and long syllables. The final ordering of the grid may have been settled only towards late Vedic times, but fundamentally the *Ṛgveda* conforms to the structure of the standard syllabary (with some deviations from the rigidity of the *mātra* in its metres). If anything, it is phonetically more elaborate; there is an extra degree of freedom: syllables were endowed with a pitch taking three values, something that can be discerned in its chanting to this day.

Oral learning puts a premium on memory power and any aid to enhancing that power is bound to have been highly valued – the transience of *śabda*, meaning in this context both 'sound' and 'word', is an insistent theme in the writings of grammarians, linguists and philosophers until very late. We do not need experiments in cognitive psychology to know that a very effective way of reinforcing auditory memory is to superimpose a pattern or structure, rhythmic or melodic or both, on the stream of sound that is to be learned 'by heart'. Because it is rigidly governed by the rules of Sanskrit prosody, versification provides such a rhythmic structure to oral text; in the context of the Vedas, it turns speech into chant and, together with the accents denoting the pitch, gives it a subliminal musicality as well. As though in tribute to the power of sound

[4]Kumud G. Ghurye (1950), *Preservation of Learned Tradition in India* (Bombay); J. F(rits) Staal, *Nambudiri Veda Recitation* (1961); both cited in [St-RM].

as the means of perpetuating what had by then come to be considered divine revelation, 'the word' in the abstract itself became a goddess, Vāc, and the metres utilised in the *saṃhitā* were attributed mystical and magical qualities (*Taittirīyasaṃhitā*); some of the Brahmanas almost deify them. The science of prosody, one of the six limbs of the Veda, was called *chandaḥ-śāstra* and *chandasi* is the qualification Panini used to denote Vedic grammatical usage – it is as if verse and Veda had coelesced into one holy unity. As in the case of grammar, a proper theoretical treatment of prosody came much after the composition of the Vedas, in the work of Pingala, and we shall have occasion to study the mathematical aspects of it later (Chapter 5.5).

But turning everything into verse was not the only device used to help fix in the mind large bodies of literary material – some of the early and middle Vedic texts also have major sections in prose. If there was no writing at the time of their composition (or if it was deliberately avoided), techniques of committing prodigious quantities of text to memory must have formed a part of education from very early times. As migration dispersed the Vedic people over time and communities divided into branches – with their own versions of what was once the common legacy – and got localised in their places of settlement, it became necessary to re-collect their store of remembered knowledge. The differences in the recensions of the *Yajurveda* for example, as well as of some of the auxiliary texts such as the *Śrautasūtra* and indeed of the *Śulbasūtra*, probably have their origin in selective recollection; memory itself became sacred and the word *smṛti*, 'that which is remembered', came to denote unquestioned knowledge. The hold of speech-based teaching and learning was so strong that memorisation and learning by rote survived the increased use of writing in historical times and remained part of traditional Indian pedagogy until very recently. Foreign travellers never ceased to be astonished by it. Here is what the Chinese pilgrim-scholar I-Tsing, who spent several years in Nalanda in the 7th century CE, has to say on memorisation and the oral culture (translation by J. Takakusu):

> The Vedas have been handed down from mouth to mouth, not transcribed on paper or leaves. In every generation there exist some intelligent Brahmans who can recite the 100,000 verses [of the Vedas]. In India there are two traditional ways by which one can attain to great intellectual power. Firstly, by repeatedly committing to memory the intellect is developed; secondly, the alphabet fixes one's ideas. By this way, after a practice of ten days or a month, a student feels his thoughts rise like a fountain, and can commit to memory whatever he has once heard (not requiring to be told twice).This is far from being a myth, for I myself have met such men.

We cannot of course tell whether there was some exaggeration in I-Tsing's account, but it cannot be doubted that the amount of textual material committed to memory was enormous. Correspondingly, we should also not be surprised at the elaborate and innovative mnemonic devices put to use, both to aid in the

memorisation and to eliminate errors of transmission. Two classes of textual variations arising out of such devices are worthy of notice as they also have a grammatical dimension: the *padapāṭha* ('word-by-word reading' as opposed to *saṃhitāpāṭha*, 'reading together') and the *prātiśākhya* ('according to each branch', branch here refers to the four Vedas in their different recensions). The grammatical point, to which we will return in the next section, concerns *sandhi*, the canonical phonetic transformations that take place at the junction of two words when they are pronounced together, generally involving the last syllable of the first word and the first syllable of the second word. In the *padapāṭha* one undoes the transformation, by breaking up a running text, say $A'B'C'D'\cdots$, where each primed letter stands for an irreducible word after it is transformed by the application of *sandhi* rules, into the sequence $A-B-C-D-\cdots$ of the pristine forms of the same words, the dash denoting a pause. (A word in the middle, say B, will undergo a change at both ends as it is joined to both A and C). The *prātiśākhya* variations carry this analysis a step further by imposing patterns of repetition. The simplest repetition turns $A-B-C-D-\cdots$ into $A'B_1-B_2C_1-C_2D_1-\cdots$ in which the *sandhi* rules are reapplied to each pairing but not between pairs – in the pairing $A'B_1$ for example B changes only in the first syllable, to B_1, and in B_2C_1 B changes only in the last syllable, to B_2, and so on. It appears complicated but, in execution, the recitation acquires an easy flow that sounds quite natural. The idea presumably was to impose an auditory structure on the text that makes it easier to remember, never mind the semantic damage and the fact that it doubles the length of the original text. The important point for us is that such mnemonic techniques (for more detail see [St-RM]) make sense only in a culture in which the human voice was the sole medium of communication. And they seem to have been successful, as we know from the survival of very long texts to the present.

To the reader who may wonder about the purpose of this excursion into the byways of spoken language, the answer is that orality influenced in a very significant sense not only how mathematics was taught and presented but also the way it was done and, even, thought about. It is of course natural that the structure a language acquires as it emerges from its formative phase will be determined partly by the medium of communication – the stress on phonetic aspects in Sanskrit is an illustration – and linguists and philologists have written on it. But we are used to think of mathematics as universal and context-free, at least in its core, and only superficially coloured by the natural language through which (a part of) it is communicated. In that perspective, it is something of a shock to see written in the great 16th century CE Malayalam text *Yuktibhāṣā* of Jyeshthadeva[5] (my italics):

[5]The only translation into English (or any other language) of the work is K. V. Sarma's ([YB-S]). It is not always as faithful to the original as one can wish. The quotations from *Yuktibhāṣā* in this book (of which there will be quite a few) are in my translation unless otherwise indicated.

> .... there is no end to the *names* of numbers; *hence* [we] cannot know [all] the numbers themselves and their order.

The sentence comes at the very beginning of the book, after it lists the *names* of the first seventeen powers of 10 as an introduction to the recursive construction of decimal numbers and their unboundedness. The immediate epistemic link is to the position of a school of linguists and philosophers, most eloquently propounded by Bhartrhari (5th or 6th century CE), that an abstract thing exists only by virtue of its having a name, an articulated sequence of syllables that defines it; the vital link to orality is obvious. The powerful hold an orally expressed Sanskrit had on the collective intellect, first exerted in early Vedic times, never weakened, even after writing became widespread.

We shall meet many examples of the close interplay between language and mathematics in what follows.

1.4 Grammar

The preoccupation with a systematic science of uttered sound did not end with a well-ordered syllabary concerned with units of sound in isolation and with how the simple consonants of the table (Introduction, section 4) combine with vowels to produce voiced syllables. I have already mentioned there the *mātra* values to be assigned to syllables preceding composite consonants. More generally, syllables undergo various transformations depending on the phonetic environment in which they find themselves. When two words are pronounced in succession, ease of vocalisation (or, maybe, just habit) imposes certain changes on the syllables at the junction, namely the last syllable of the first word and the first syllable of the succeeding word. These are regulated by the rules of *sandhi* (junction) which are purely phonetic rules having nothing to do with the meanings of the two words. The *padapāṭha* and the *prātiśākhya* referred to in the last section provide an empirical data base of how these rules operated in the Vedas, but we have to wait until Panini, six or seven centuries after the earliest Vedas, for the first account of their formalisation.

The roughly 4,000 *sūtra*s of the "Eight Chapters" of Panini, *Aṣṭādhyāyī* (6th or 5th century BCE), did not just classify the junctions into a finite set of categories, each category characterised by a precise rule of formation that operates within it, determined by the syllabic context specific to it. Similarly general methods of categorisation were applied to every component of a linguistic whole at its various levels of organisation, from the sentence down to the syllable (sometimes with explicit exemptions; Panini was after all dealing with a living language). That makes for a highly formal work, which is another way of saying that it gives priority to the syntactic aspects of Sanskrit over the semantic. It has been said that every sentence constructed following the Paninian rules will be grammatically correct and, conversely, that every grammatically correct sentence will be found in a (theoretical or imaginary) compendium of all Paninian sentences; in other words, the totality of the *sūtra*s form a complete

and self-consistent framework encompassing all grammatically legitimate linguistic expressions. It is also extremely economical. Sanskrit, with its emphasis on sounds and their production and the contextual transformations they are subjected to, lends itself perfectly to such a treatment. Modern linguists who, consciously or otherwise, follow in Panini's footsteps in their analysis of 'language' (in particular European languages with their haphazard organisation of sound patterns as exemplified in the randomness of their alphabets) can achieve nothing like that level of completeness and generality, not to mention economy.

The comprehensiveness of Panini's analysis of the structure of Sanskrit is matched by the generality of the theoretical machinery that he created for the purpose. The methodological framework that results is, necessarily, complex and subtle. A schematic (and grossly inadequate) outline of the organisational principles on which it is based will start with the words that constitute a meanigful sentence as classified by their functional role (the 'parts of speech'). They are then analysed into smaller, more primitive or elementary, units (like atoms, to use the simile of the much later Bhartrhari) such as noun stems, verb roots, prefixes and suffixes, etc. Each of these general classes can be categorised further, e.g., nouns by gender, number and case, verbs by tense and mood, and so on. The ideal rules which operate on the elements will be category-specific, uniform for each member of a category, and such operations may leave the object(s) on which they operate in the same category or take them to another category. For instance, conjunction of nouns leads to another noun, as in the additon of numbers as an operation on number names: *eka* (1) and *daśa* (10) is *ekādaśa* (11), the junction rule causing the lengthening of the last syllable of *eka*. On the other hand, *dvirdaśa* (20) results from the two-fold repetition of the action of counting to 10, i.e., from an operation on the pair *daśa* (which in this example is an adjective qualifying a noun, say cows) and *dviḥ* (numerical adverb which is a recognised subcategory), leading to another adjective. In contrast, the formation of *dvādaśa* (12) from the same names *dvi* and *daśa* follows a different rule, similar to the one which applies to *ekādaśa*.

If the reader suspects from all this that Panini was applying abstract set-theoretic ideas to the study of grammar, he or she will be entirely right. Such an approach would not have been feasible or productive without the invention of several metalinguistic (metamathematical) constructs. Sets and operations on them are categorised by the use of special notation or (metalinguistic) markers (which are to be removed in the final readable text) subject to (meta) rules, just as there are metagrammatical rules governing the operation of the grammatical rules; the formation of words denoting repeated action, for example, falls under the marker *suc*. The metarules themselves are subject to a hierarchical order, so as to prevent them from clashing: when two rules lead to contradictory constructions, it is specified which one of them overrides the other. The inescapable conclusion is that Panini utilised, in the context of Sanskrit grammar, a very effective tool of modern mathematical reasoning: to think of operations as maps from sets to sets, $\{\text{nouns}\} \times \{\text{nouns}\} \to \{\text{nouns}\}$ (i.e., composition within a set as we think of it today) and $\{\text{adverbs}\} \times \{\text{adjectives}\} \to \{\text{adjectives}\}$ in the two

numerical examples above. It is to be noted that this point of view naturally accommodates an older Indian concern, the idea of recursion – though Panini does not explicitly mention it – since the output of a composition can be fed into a subsequent composition. He also understood another and deeper mathematical truth, that the path to generalisation passes through abstraction. It is no wonder then that the *Aṣṭādhyāyī* continues to hold an unending fascination not only for grammarians and linguists, but also for logicians, mathematicians and computer theorists, in short for anyone interested in formal structures.

There are two levels, which we may call the concrete and the abstract, at which Panini's work impinged (or ought to have impinged, see below) on the practice of mathematics in India. The literary style of presenting mathematics made it obligatory for any writer to be well-versed in the intricacies of Sanskrit grammar, whether he wrote in verse or prose. This is a relatively undemanding accomplishment; in fact a traditional education in many parts of India, even in recent times, began with Sanskrit and its grammar (*vyākaraṇa*) before specialisation in a chosen science or *śāstra*. More important at this concrete level is the skill to construct a precise technical vocabulary as and when needed, especially in the quantitatively formulated sciences such as mathematics and astronomy – words once formed cannot leave room for ambiguity when analysed for their exact meaning. As a general rule, this requirement is seen to have been met in the mathematical texts, once we make allowance for the role of usage and consensus in the simplification of widely used terms. Thus Aryabhata used the term "half-chord" (*ardhajyā* or *jyārdha*) for the sine of an arc of the unit circle, which is actually half the chord of twice the arc – confusingly, he also used the same term for the difference of two sines – and that quickly got further shortened to just *jyā*, chord.

The fundamental purpose of creating an unambiguous terminology is obviously semantic – does a particular expression mean exactly what it is intended to? – though the means adopted may be syntactical. The standard method in Sanskrit of making new words from old is nominal composition, predominantly through the process called *samāsa* and variations thereof. The structure of Sanskrit allows the telescoping of words, essentially without bound on the number of words so joined together, for example nouns with nouns or with adjectives. Thus, in the phrase 'three hundred and sixty days', all the words can be telescoped into one and the process will involve several stages of word-compounding: first, 'three hundred and sixty' as one unit, qualifying the noun 'days', then 'three hundred' and 'sixty' as independent units conjoined (additive), then 'sixty' as derived from'six' and 'ten' and 'three hundred' from 'three' and 'hundred' (both of them multiplicative or repetetive). This particular example occurs in the *Ṛgveda* in an archaic form and the telescoping there is only partial.

There is of course a typology of nominal composition and there are rules of composition for each type. Since the process operates at the level of words, the classification and nomenclature of types are primarily based on semantic considerations. But when words are brought together, the junction rules must

come into play, giving them a strong role in the implementation of the composition. The occasional difficulty arises in the *decomposition* of a given compound word into its component parts so as to get at exactly what was meant to be conveyed. Generally, but not always, the context will provide a clue. I give two examples of not-so-rare pitfalls from the domain of mathematics. Western scholars did not see the text of *Āryabhaṭīya* before the publication of Kern's edition in 1874, but had known of its existence from Brahmagupta's reference to it as *Āryāṣṭaśata*. Two of the 19th century (pre-Kern) pioneers, Colebrooke and Burgess, took *aṣṭaśata* to mean 800, whereas the correct meaning, factually and grammatically (we shall see why later), is 108, for the number of verses in the non-introductory portion of the book. The second example concerns the term *tricaturbhuja* in Brahmagupta's work on cyclic quadrilaterals: how can a rectilinear figure be both "three-armed" and "four-armed" (or "three-four-armed") at the same time? Plausible explanations have come only very recently; the term does denote a cyclic quadrilateral. One explanation[6] decomposes it as a particular *samāsa*, according to which it is a quadrilateral which shares with all trilaterals a particular property (that of having all its vertices on a circle). Another, less forced, relies on an analysis encoding the fact that a cyclic quadrilateral is really three-sided because it has only three independent sides (see Chapter 8.4 below).

At the abstract level, the impact of Panini's structural and formal treatment of grammar on the way mathematics was done, as opposed to how it was 'written up', is more problematic. A century or two (or three) after Panini, Pingala approached the problem of analysing and classifying all possible metres in Sanskrit verse, subject to given constraints (on the number of syllables in a line and, sometimes, the number of lines in a verse), in a very Paninian spirit. Apart from posing and answering the combinatorial questions that arise – for example: how many distinct sequences of a given length can be constructed from a two-valued variable? – the work is notable for the structural point of view adopted throughout. One of the methods employed is the following: first classify all ordered sequences of a small number (three) of syllables (called a *gaṇa*), denote each such sequence by an abstract symbol (a Sanskrit syllable functioning as a metalinguistic marker) and describe the structure of the line as an ordered sequence of such markers or set labels. But Pingala is an exception; by and large, mathematicians seem to have been content with posing and solving problems in isolation as they emerged in particular circumstances, each by its own special method. It is rare to see attempts at classification of mathematical objects into sets with a defined structure, whose members have some property in common and on which similar or identical operations can be performed. Though the idea of denoting unknown numerical quantities by 'letters' (syllables) and doing arithmetic with them as though they were determinate numbers dates back to Brahmagupta or the Bakhshali manuscript

[6] Pierre-Sylvain Filliozat, "Modes of Creation of a Technical Vocabulary: the Case of Sanskrit Mathematics", Ganita Bharati, vol. **32** (2010) p. 37.

(whichever is earlier), it is striking that there was no attempt to characterise them as belonging to structurally defined sets: integers, rationals or reals. In the 12th century, Bhaskara II got round to writing a tract *Bījagaṇita*, “Algebra”, devoted to methods of solution (including Diophantine solutions) of algebraic equations, but it has no explicit statements about the class of numbers within which a solution is sought. Later on, we will come across several instances in which this absence of an abstract, Paninian mindset might have acted as a brake on natural (and, to the modern eye, obvious) generalisations. With the partial exception of Pingala, it is only in the 15th century that we finally glimpse the Paninian paradigm beginning to have an influence in mathematics; *Yuktibhāṣā* has a passage which, somewhat uncertainly, defines positive integers by the property of succession and it also has a long discussion of polynomials (and rational functions) as a category. These topics will occupy us later.

The *Aṣṭādhyāyī* is a text like no other; it is in fact best thought of as a metatext, written in a metalanguage, using Sanskrit syllables and words which serve specific technical ends and which often do not have their normal meanings. There is a section called *paribhāṣā* to supply the key to the translation of the *sūtra*s into ordinary language. An analogy will be a text on logic written entirely in formal logical notation and equipped with a glossary enabling the conversion of symbolically expressed propositions into, say, plain English. Not surprisingly, it has engendered a vast lineage of commentaries – of which the most insightful is still, probably, the first one that has survived: Patanjali's *Mahābhāṣya* – but none deviated from the principles set down in it. Perhaps out of admiration for its perfection (or intimidated by it), no author followed its metatextual format either. In these respects, the closest parallel is with *Āryabhaṭīya* which also inspired many commentaries and explanations over a long period of time, few of which brought significant generalisations to Aryabhatan doctrine until we come to Madhava.

A final remark on the *Aṣṭādhyāyī* as it relates to this book. Just as the grammatical principles expounded in it continued to govern Sanskrit far into the future, so did, antichronologically, Vedic grammatical practice from 700 years earlier also conform to it, to a surprising extent. It is this fortunate circumstance, aided by its classification of exceptions, mostly to accommodate Vedic usage, that allows an unerring decryption of Vedic number names by means of the rules of the *Aṣṭādhyāyī*. In the absence of a written notation, the nomenclature of numbers becomes our only means of understanding the development of decimal place-value enumeration in India. That involves an analysis of nominal composition, fairly straightforward most of the time. The occasional ambiguities which resist Panini can generally be resolved by an appeal to the *padapāṭha*. In any case, the importance of Paninian analysis of number names cannot be exaggerated; it allows us to assert that the principles of decimal enumeration were established before the poems of the *Ṛgveda* were composed.

2

Vedic Geometry

2.1 The *Śulbasūtra*

It may come as a surprise that I am going to invert the historical order and take up Vedic geometry before turning to the geometrical knowledge and culture of the Indus Valley people. The reason why this is the practical thing to do is in the nature of the sources. The state of geometric knowledge during the Vedic times is documented and authenticated by the *Śulbasūtra*, even if they provide little justification for the constructions, or substantiation for the statements made. A portrait of the mathematical culture of the Harappans has, in contrast, to be reconstructed from the indirect evidence of artefacts and will therefore remain more or less conjectural. Every bit of help we can get to make the picture a little clearer is to be welcomed; and looking at Harappa through the lens of the geometry of the Vedic texts does clarify some points and perhaps even hints at possible continuities.

The *Śulbasūtra* are texts concerned with the architecture of the Vedic ritual arena, including the geometric forms and measurements of the various elements of it. The floor plans are also, correspondingly, plane figures of different shapes, generally with a high degree of symmetry, even those that are quite complicated. The texts describe the architecture not only in terms of the actual building – the shapes and sizes of the bricks to be used and how to put them together, for instance – but also more abstractly, in geometrical terms. It is the geometry alone, together with its mensurational setting, that will concern us here. The presentation of the geometric material takes several forms: constructions of figures of desired shapes and dimensions; magnifying the area of a figure by a given number by scaling its linear dimension; transforming a figure into another while preserving the area; statements, either explicit or conveyed implicitly, of general geometric truths ('theorems'); etc.

The origins of the preoccupation with geometry in the context of ritual activities and, for that matter, of Vedic rituals themselves, are lost in antiquity. Leaving aside for the time being the question of a possible link with

P. P. Divakaran, *The Mathematics of India*, Sources and Studies in the History of Mathematics and Physical Sciences, https://doi.org/10.1007/978-981-13-1774-3_2

Indus geometry, the earliest texts, especially the *Yajurveda* (the Veda of rituals), already have many references to rituals and the associated altars. The gathering together of specialised architectural knowhow in the *Śulbasūtra* happened later, between the 8th and the 4th centuries BCE, as part of the formal setting down of religious practices in the texts called the *Śrautasūtra* and of the composition of the Vedanga texts. Several versions of the *Śulbasūtra* are known to have existed, through references in other texts or through hearsay. Of these, four have survived, presumed complete and intact, named after Baudhayana (Baudhāyana), Apastamba (Āpastamba), Manu or Manava (Mānava) and Katyayana (Kātyāyana); the chronological order is most probably the order in which the names are given above, though Manava is somewhat difficult to place precisely. In spite of the long time frame of their writing, the mathematical content is quite similar though the style of presentation and the order of the topics vary from version to version. It is certain that they represent different recollections of a common store of geometric and architectural knowledge from an earlier epoch; the existence of many recensions of roughly the same material has, very likely, to do with social or clan differentiation and geographical and temporal patterns of migration.[1]

The word *śulba* (or *sulva*) means a cord or rope and *sūtra* describes the style of the writing, a compressed aphoristic presentation in which all inessential words (including, often, the verbs) are dropped. The geometry described in these "Cord Sutras" is the geometry that can be done with (inextensible) cords. The two basic elements out of which the geometry is built up are therefore the circle (produced by fixing one end and taking the other end of a stretched cord round) and the straight line (produced by stretching the cord end to end; the use of a bamboo straight-edge is also common). In other words, they deal with 'compass and ruler' geometry. Measurements and numbers occur freely, but basic geometric truths when they are invoked are generally metric-free. Thus the two fundamental principles on which all *Śulbasūtra* geometry is based can be stated as:

1. The circle is the locus of points which are equidistant from a fixed point, its centre.

2. The straight line is the shortest distance between two points; or, in the *Śulbasūtra* spirit: two fixed points along a curve are farthest apart when the curve is a straight line (i.e., when the cord that is the curve is stretched taut).

The geometrical point is not explicitly mentioned anywhere; the notion occurs frequently both as intersections of arcs and straight lines and as their constituent parts, but was not considered to need further explanation. If Baudhayana were Euclid, he might have thought of the two principles defining circles and stright lines as postulates. More likely is that he looked on them as the

[1] My account of *Śulbasūtra* geometry owes a great deal to the critical edition of the four texts, with English translation and commentary, by Sen and Bag [SuSu-SB]. Sarasvati Amma's book [SA], whose first version was prepared in the 1950s (and hence without the benefit of having [SuSu-SB] to hand) is also excellent; the only disppointment is its rather cursory treatment of the transformation between the square and the circle.

elements from which more complex geometrical objects were to be built up about which, in turn, valid geometrical statements could be made and verified.

A last general observation that will help provide a perspective is that the geometrical quantity that is of greatest interest in these texts is the area of plane figures. The Sanskrit name for a bounded part of the plane is *kṣetra* ('field'), also used to denote its area, and geometry came to be designated as *kṣetragaṇita*, the mathematics of planar figures and their areas.

Because the plans served an architectural purpose, they have to be thought of as being made on a level piece of ground. Going by the later literature, horizontality was ensured by means of a water level and the vertical determined by means of a plumb line. The ritual enclosure itself and all structures within it had the east-west line (called *prācī*) as the primary axis of symmetry, a convention which had its distant origin in Indus Valley town planning and architecture and came to be the default choice of alignment in future sacred architecture (with many deviations imposed by competing considerations and local conditions). In geometry itself, east became the invariable cardinal direction, the positive y-axis, in all future work, until the very end. The two later *Śulbasūtra*, those of Katyayana and Manava, have a procedure, the same in both the texts, for finding the true east by the use of shadow-geometry. I describe it briefly because it is the prototypical illustration of the original Indian view of the connection between circles and orthogonal pairs of lines, already present in Indus Valley geometry.

Fix a gnomon (here a straight rod vertically driven into level ground) of suitable height at the chosen point and draw a circle of a suitable radius with that point as the centre. As the sun traverses the sky, the shadow of the gnomon will change in direction and length and its tip will fall on the circle twice during a day, once before noon at the point W and once after noon at the point E. Then the directed line EW marks the east-west direction. Figure 2.1 shows the geometry (as seen from above) when the sun is to the south of the location of the circle. Once the east-west axis is established, the other cardinal directions are found by using one of several different methods (see the next section) for drawing a perpendicular to a given line.

How true the east determined by this method is depends of course on how 'east' was defined. If the east-west line is defined to be in the plane passing

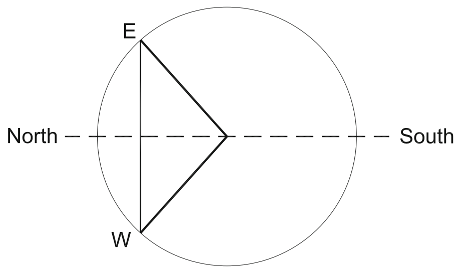

Figure 2.1: Shadow-determination of the true east (*prācī*)

through the centre of the circle and perpendicular to the axis of the earth's rotation (the plane of the apparent daily motion of the sun) in any epoch, then the shadow method gives the 'true' east at that epoch, independent of latitude and the day of the year or season. If 'east' is defined by reference to some celestial coordinate system, then its direction will change with changes in the orientation of the earth's axis with respect to the fixed stars. In the absence of any evidence of precise position measurements of individual stars in the Vedanga, it is best to take the *Śulbasūtra* method as operationally defining the east. In particular, if the location is north of the tropic of cancer (which is the case for the place(s) of composition of the *Śulbasūtra*), the sun will never rise due east of it. These are small points but they are of some relevance to efforts at interpreting archaeological data from prehistoric times by astronomical considerations.

We may also wonder how errors in horizontality of the ground and verticality of the gnomon affect the accuracy of the method. It is a simple exercise to repeat the Vedic procedure and to conclude that a precision of 1 degree or less is easily attainable. There are minute 'theoretical' errors as well, for example, the sun's latitude will change slightly during the transit of the tip of the shadow from W to E.

2.2 The Theorem of the Diagonal

The theorem of the diagonal is the geometric result commonly called Pythagoras' theorem. It has been known (outside the circle of traditional Indian astronomers, that is) since the 1870s at least that the statement of the theorem occurs in the earliest *Śulbasūtra*, that of Baudhayana (as well as in others), and that Baudhayana lived around the 8th century BCE. The *Śulbasūtra* do not give it a name, but later texts often refer to it by names which are well captured by the heading of this section. The first occurrences of the result formulate it as applying primarily to the two adjacent sides and the diagonal of a rectangle but, in the *Śulbasūtra* already, it is used to construct similar right triangles into which more complicated altar shapes are decomposed, as a means of scaling up their areas. Nevertheless, the nomenclature stuck forever, even when the figure considered was the right triangle on its own. For instance, the Nila texts use the expression *bhujā-koṭi-karṇa-nyāya* which translates as 'the principle of the base, the altitude and the diagonal', with the tacit understanding that *karṇa* means the hypotenuse in the context.

Chapter 1, *sūtra* 12 of Baudhayana, BS 1.12 in short, says (in the translation of Sen and Bag [SuSu-SB] as are all direct quotes from the *Śulbasūtra*):

> The areas (of the squares) produced separately by the length and the breadth of a rectangle together equal the area (of the square) produced by the diagonal.

This is followed by BS 1.13:

> This is observed in rectangles having sides 3 and 4, 12 and 5, 15 and 8, 7 and 24, 12 and 35, 15 and 36.

Just before this (in BS 1.9), we have the special case:

> The diagonal of a square produces double the area (of the square).

BS 1.10 and 1.11 then explain how to obtain a square whose area is three times that of the original square. Apastamba (1.4) has both the general theorem and the special case (phrased differently), but stated in the reverse order – it is as if what was a process of induction for Baudhayana, from the special case to its generalisation, became a passage to the corollary for Apastamba. Manava's statement has the same content but has a mathematical twist that is sufficiently intriguing for us to return to it later.

In looking at these statements and assessing their historical significance, it is best to distinguish between the geometric theorem (as in BS 1.9 and 1.12) and the arithmetical question of the existence of diagonal (or Pythagorean) triples of integers or rational numbers $(c; a, b)$ satisfying the equation

$$c^2 = a^2 + b^2$$

as in BS 1.13; until a connection is made, these two phenomena, one geometric and the other arithmetical, need have nothing in common. The interplay between geometry and arithmetic or number theory of the kind exemplified by the two avatars of the diagonal problem has, as we know, always been close – we shall meet other examples from India as we go along – and is a deep and very fertile branch of modern mathematics.

To begin with the arithmetical side first, it is useful to note that if $(c; a, b)$ is a diagonal triple, so is $(rc; ra, rb)$ for any rational r. In particular, mutiplication by a common denominator of a, b, c converts a fractional triple into an integral triple. So, in principle, it is not necessary to distinguish between the rational and the integral cases and a, b and c can be resricted to (positive) integers in the present discussion. The scaling property was known to the *Śulbasūtra* authors: Apastamba (in 1.2, before the enunciation of the geometric theorem), has a procedure for enlarging areas of triangles that depends on scaling and, most explicitly, Manava (11.17) says:

> The sides (of a right triangle) are made with 3, 4 and 5. Those of others are made by multiplying (these numbers) with desired (quantities) as they may be required in the (construction of) altars; this has always been prescribed by ancient teachers.

Who were these "ancient teachers" to whom we owe the idea of similar triangles: Baudhayana and Apastamba or even more ancient ritual experts?

Sets of diagonal triples, both integral and fractional are found in all the *Śulbasūtras*. Many of them are related by scaling but the texts make no distinction between integral triples which have no common factor and those which are obtained by scaling from known triples. That may not have been an important consideration in the making of altars; in fact, much of the geometry

is concerned precisely with scaling, in part because the dimensions of the altar were determined by the height of the patron or sponsor of the particular performance of the ritual.

The lists of diagonal triples and the use of scaling raise the question of whether the *Śulbasūtra* authors knew that the sequence of triples is unending and whether they had a general method of finding the primitive integral triples, i,e., those in which a, b and c have no common factor. Sarasvati Amma ([SA]), following traditional Indian commentators of Apastamba, suggests that they did know how to construct one particular infinite sequence. To see why and how, let us start with a standard characterisation of primitive triples and work our way back to what is in the texts (in actual fact, two *sūtra*s in one particular text). All primitive triples are included in the formulae (a result going back to Euclid and also stated by Brahmagupta in his study of rational right triangles and (cyclic) quadrilaterals, see Chapter 8.4)

$$c = m^2 + n^2 \qquad a = m^2 - n^2 \qquad b = 2mn,$$

where m and n are (positive) integers with $m > n$. m and n can be further restricted since, if they are both odd or both even, a, b and c will all be even and the triple will not be primitive. For $m = n + 1$ in particulr, we get

$$c = 2n^2 + 2n + 1 \qquad a = 2n + 1 \qquad b = 2n(n + 1),$$

or, by dividing by $2n(n+1)$, an infinite sequence of rational right triangles with one of the short sides having length 1. After a simplification, we get

$$c = 1 + \frac{1}{2n(n+1)} \qquad a = \frac{1}{n} - \frac{1}{2n(n+1)} \qquad b = 1.$$

For $n = 1$ for example, the triple is (5/4; 3/4, 1).

The connection of these formulae to certain prescriptions in the *Śulbasūtra* comes about through the use of diagonal triples to draw right triangles *via* the converse of the theorem of the diagonal. One of the ways this was done is as follows. For any given rational triple $(c; a, b)$, draw on the ground a line AB of length equal to b, take a cord of length $a + c$ and fix its ends to the ground at A and B by means of pegs after making a knot at a distance a from one end, say A; pass another peg through the knot and drive it into the ground when the cord is stretched taut by pulling on this peg. This fixes the point C so as to make ABC a right triangle with the right angle at A.

This procedure and the formulae for a, b and c in terms of n, together, provide a simple and practical way of generating an infinite sequence of rational right triangles. Since $a + c = 1 + 1/n$, by increasing the length of the cord by a fraction $1/n$ of its original length and placing the knot so as to increase the length of the hypotenuse by a fraction $1/(2n(n+1))$, we get a new rational right triangle for each $n = 2, 3, \cdots$. The question is whether there is any evidence that the general formula was actually used. What is in the text (AS 1.2) is an explicit mention of the case $n = 2$, for which $a + c = 1 + 1/2$, $c = 1 + 1/12$.

Whether the general case was understood turns on the reading of the next *sūtra*, in fact on the precise significance of a particular prefix. Later commentators (who are from the medieval period) say that the general case was covered. The default option is to take commentators seriously; but instances in which they ascribed their own insights to the original texts are not unknown (the commentators, at least, knew the general formula). It is also true that texts often follow the practice of dealing carefully with a simple special case first and disposing of generalisations with a terse remark (such as, possibly, here or in Manava's remark on scaling the (5; 3, 4) triangle, quoted above). What cannot be doubted is that the *Śulbasūtra* sages were thoroughly familiar with rational diagonal triples.

The observation i) of a finite number of primitive triples and ii) that they can be scaled by rational numbers to generate an infinite number of rational triples takes us part of the way to the general geometric statement of the theorem of the diagonal (as given in Baudhayana 1.12 and Apastamba 1.4), including the special case of the square. Only part of the way because, as we can recognise easily today, to make the deduction logically complete one needs to start with all primitive triples and extend the scaling to all real numbers. There is no evidence at all that the authors of the *Śulbasūtra* had an appreciation of real numbers which are not rational even though $\sqrt{2}$ is a constant presence in the texts, through the diagonal of the square. (We have to wait till Nilakantha and the 16th century to see this distinction made, as will be seen later). We have also no evidence that there was an understanding of the need, or even the idea, of starting with all primitive triples. The question is of importance in the context of the fact that scholars, traditional and modern, have written at length about what the original justification (proof) of the geometric theorem might have been, whether it was arrived at inductively by staring hard at enough examples of rational triples.

For the special case of the square, the general consensus is that there were geometric proofs based on a method of matching areas, perhaps because it is easy to visualise such proofs. Two variants that have been suggested, both relying on the way the result is used later on in more complicated constructions, are shown in Figure 2.2. In both figures, d is the diagonal of a square of side a. In the first, the square on d is divided into eight congruent right triangles, four of which form the square on a. In the second, the square on d is cut into four congruent triangles of which two can be fitted by 'cutting and moving' into the square on a.

In contrast to the special case, the question of what arguments the *Śulbasūtra* authors had for the general theorem has remained a matter of great uncertainty. In the absence of even a remote hint about their thinking, we can only guess (ignoring Sarasvati Amma's ([SA]) injunction: "To speculate . . . is idle"). There are two possible routes they could have followed. The more popular view used to be that it was just a brilliant guess based solely on the examples of diagonal triples listed in the texts. That is a big extrapolation but it gains in credibility if we take into account the scaling property that Manava took note

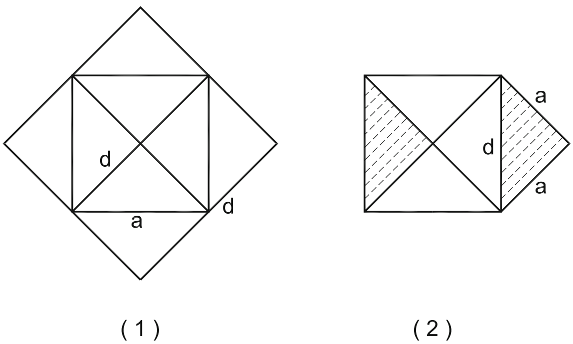

Figure 2.2: Two area-matching proofs of the theorem of the diagonal for a square.

of and Apastamba's modification of the basic cord construction to generate new triangles which are not related by scaling. On the other hand, neither of these lines of generalisation leads to the special case of the isosceles right triangle as its hypotenuse is irrational; so it is entirely natural to wonder whether a direct generalisation to the rectangle of the geometric proof that worked for the square may not have been sought by the authors of the *Śulbasūtra*. The texts contain enough evidence of their ingenuity in manipulating areas, quite a few of them involving the diagonal theorem, for us not to be too hasty in discarding the possibility.

I give below a small sample of proofs of the general theorem of the diagonal, chosen from across the post-*Śulbasūtra* history of geometry in India, each of them with its own distinctive underlying idea – whose possible origins I have also tried to indicate – and with a different emphasis, including one that is quite Euclidean in spirit. But before doing so, it is pertinent to remember the central role the theorem played throughout this long history, culminating in the infinitesimal geometry of the calculus of the Nila school. Given this preeminence, it should not be surprising that ways of proving it also served as a model topic in Indian discussions of the epistemology of mathematics. For instance, around 1525 CE, Nilakantha, the great astronomer-mathematician-polymath of the Nila school, produced an analytic commentary (*vyākhyā*) on his own work called *Siddhāntadarpaṇa* (adequately translated as "The Mirror of the Doctrine") in which he summarised the state of astronomical knowledge of his time including his own contributions to it. Despite its brevity – the original and the commentary together take up about 30 printed pages – it is a mirror held to the questions he considered the most significant in the Indian view of epistemic foundations. Among other things, he sets down his views on acceptable methods of acquisition and validation of astronomical and mathematical knowledge and brings up the theorem of the diagonal for the square as a result that should be "self-evident to the intelligent". He then goes on to give yet another proof of the general theorem. It is as though the fascination of

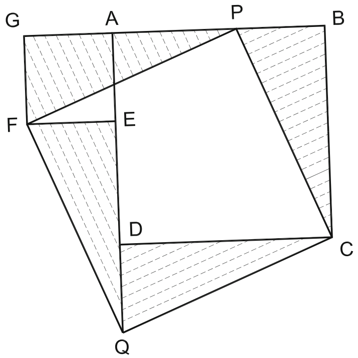

Figure 2.3: The *Yuktibhāṣā* proof

this miraculously powerful result never faded. Nilakantha himeslf calls it "the fundamental basis of geometry".

The first proof I describe is a cut-and-move proof from *Yuktibhāṣā*, only slightly different from (and contemporaneous with) Nilakantha's. Draw two squares of sides a and b as shown in Figure 2.3 (squares $ABCD$ and $GAEF$ respectively). Mark P on AB and Q on AD extended, such that the lengths of BP and DQ are b (so length $GP = a$). Then we have four congruent right triangles CDQ, CBP, PGF and QEF with no overlap. If c is the length of the hypotenuse of the triangles, $FPCQ$ is a square of area c^2. Triangles CDQ and QEF lie entirely within the square on c and PGF and CBP lie entirely within 'the sum of the squares on a and b'. Now match the triangles pairwise as indicated by the shading of the figure. The unshaded part is common to the square on the diagonal and the sum of the squares on the sides. That completes the proof. The instructions in *Yuktibhāṣā* about how to manipulate these triangles by cutting and moving are quite graphic and very much in the *Śulbasūtra* spirit; there is no need to know the formulae for areas for example.

A second proof combines area matching with the formula for the area of a right triangle and so has an algebraic flavour. $ABCD$ is the square (of side c) on the hypotenuse AB of the right triangle APB of sides a and b with $a > b$. By duplicating the construction on the other three sides of $ABCD$ as hypotenuse, we get the diagram of Figure 2.4(1). By symmetry (or by Euclidean matching of lengths and/or angles), the triangles ABP, BCQ, CDR and DAS have the same area $ab/2$ (they are in fact congruent) and $PQRS$ is a square of side $a-b$. Hence

$$c^2 = 4 \times \frac{1}{2}ab + (a-b)^2 = a^2 + b^2.$$

This proof is attributed to Bhaskara II (by two of his commentators), as implied by a verse of *Bījagaṇita* stating that the square of the difference of the two short sides of a right triangle together with twice their product is the sum of their squares, "as in the case of numbers", the formula used above. It is preceded by an example in which the student is asked to compute the

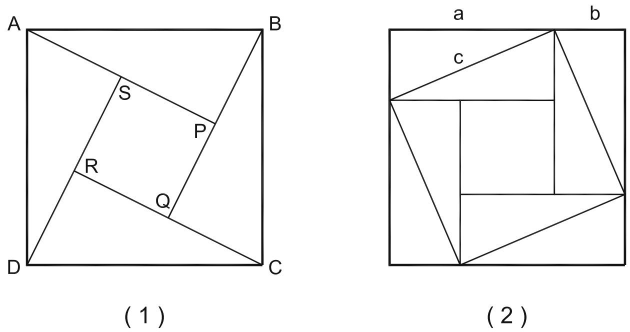

Figure 2.4: Two geometric-algebraic proofs

hypotenuse of a right triangle of sides 15 and 20 and to provide its justification. Bhaskara II does not enunciate the theorem of the diagonal in *Bījagaṇita*. He has no reason to, in a book on algebra; perhaps he only wanted to illustrate the power of algebraic methods by using them to prove a standard geometric result.

It is interesting to speculate whether a proof like this could have been found at the time of the *Śulbasūtra*. The algebraic identity used above, like all homogeneous quadratic identities in two variables, is easily turned into a statement about adding and subtracting areas of squares and rectangles. It is equally easy to establish the identity by matching areas as in Figure 2.5: $ABCD$ is a square of side a and points E, F, G, H are marked as shown, by subtracting b from each side, so that $EPHD$ is a square of side $a-b$. We can now add areas: $(a-b)^2 + 2ab = EPHD + ABFE + GBCH = ABCD + GBFP = a^2 + b^2$.

There can be absolutely no doubt that this type of demonstration was entirely within the conceptual capability of the authors of *Śulbasūtra*, as attested by several examples of the same general type (see the next section). But we have to be careful. Algebraically, from our vantage point, all quadratic identities can be generated from, say, the formula for $(a+b)^2$ by substitution $(b \to -b)$

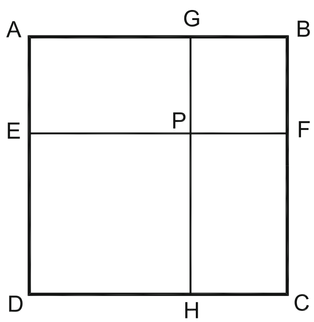

Figure 2.5: Geometric proof of the identity $(a-b)^2 + 2ab = a^2 + b^2$

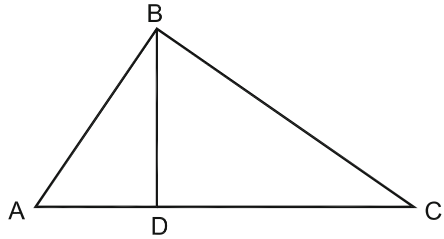

Figure 2.6: A proof based on similar triangles

and by reordering terms, using algebraic rules whose recorded first formulation is not earlier than Brahmagupta. I think it is safe to say that the savants of the 8th century BCE would not have thought of the steps involved in manipulating areas from such an abstract point of view. Consequently, they had to, and did, make a fresh geometric construction for every 'area identity' they were interested in, even when the corresponding algebraic identity is essentially the same.

There is a variant of the proof of Figure 2.4(1) that is based on the identity $(a+b)^2 = (a-b)^2+4ab$ (no transposing of terms!) as shown in Figure 2.4(2). It is self-explanatory and also illustrates the point made in the preceding paragraph..

It will of course be natural for the reader to compare these proofs with Euclid's proof of Pythagoras' theorem (Book I, proposition 47), with its construction of the squares on the three sides drawn on the outside, not overlapping with the triangle. We can wonder how much simpler the proof might have been for him if the square on the hypotenuse had been drawn inwardly so as to enclose the triangle.

Our third proof is also attributed to Baskara II by his commentators. Let ABC be a right triangle with the right angle at B and let BD be the altitude (Figure 2.6). Then triangles ADB and BDC are both similar to ABC; hence

$$\frac{AB}{AD} = \frac{AC}{AB}, \qquad \frac{BC}{DC} = \frac{AC}{BC}.$$

It follows that

$$AC = AD + DC = \frac{AB^2 + BC^2}{AC}.$$

To see how far back we can reasonably trace the idea of this proof, note that, from the similarity of the triangles ADB and BDC, we also get the result $BD^2 = AD.DC$. If we take into account the fact that the circle with AC as diameter passes through B, this last statement becomes the theorem that a diameter intersecting a chord perpendicularly divides it into equal halves with the square of the half-chord equal to the product of the two segments of the diameter. In this form, the result is in *Āryabhaṭīya*; in fact it is the second half of the verse of which the first half is the enunciation of the theorem of the diagonal (stanza 17 of the *Gaṇita* chapter). This stanza is thus the initial

step in Aryabhata's philosophy of trigonometry through the consideration of half-chords. Aryabhata of course gives no proof, so we do not know whether he used the notion of similar triangles, either for the theorem of the diameter and the chord or for other results, some of which were later proved (for example in Kerala) by techniques relying critically on the idea of similarity. The chord-diameter relationship was first noted earlier in the writings of Umasvati (Umāsvāti), a commentator on Jaina astronomy/cosmography, who lived around the beginning of the common era. It is also interesting that this proof is a special case of Euclid's proof of a theorem (Book VI, proposition 31) on similar figures on the sides of a right triangle, which itself is a generalisation of the theorem of the diagonal.

The notion of similarity of triangles is not explicitly mentioned in the *Śulbasūtra*, but it is present in the operation of scaling, which is just a way of creating a family of triangles each having the same ratio between any two sides. Apart from the statement of the principle as in the quotation from Manava earlier in this section, it is also used in the prescription for the dimensions of (right-triangular) bricks in the construction of complicated altars; one such prescription uses the fact that two right triangles ABC and $AB'C'$ in which A, B, B' and A, C, C' are collinear and BB' and CC' are parallel are similar. By the time of *Āryabhaṭīya*, it had become clearly understood that two right triangles having one pair of acute angles equal obey the 'rule of three', i.e., have their sides (picked, say, in increasing order) in constant proportion, and this principle (*trairāśikam*) became the most ubiquitous proof device in geometry; more strikingly, deviations from it became the starting point of Aryabhata's discrete calculus on the circle (see Chapter 7.5). As *Yuktibhāṣā* says, "Almost all mathematical procedures (*gaṇitakriyā*) are pervaded by the rule of three and the theorem of the diagonal". Naturally, there is no hint of any such general understanding in the *Śulbasūtra*.

I conclude this section with three remarks. Firstly, one of the legacies of the diagonal theorem is that Indian geometry adopted perpendicularity, not parallelism as in Greek geometry, as the most fundamental relationship that a pair of lines can have. A line parallel to a given line, when needed, is obtained by constructing the perpendicular to the perpendicular and the presence of parallels in a geometric situation is rarely if ever noted or made direct use of. Even the tangent to the circle at a point was thought of as the normal to the radius at that point. This preference for normality over tangency may well have acted as a constraint on the imagination of Indian geometers.

The second observation is about the interplay between the diagonal theorem and algebraic identities among homogeneous quadratic expressions in two variables. There are of course no symbols or rules for manipulating them, but the equality between the product of a and b and the area of a rectangle of sides a and b runs through many of the constructions; it is in fact the fundamental principle of what is sometimes called geometric algebra. And it goes deeper. For example, though Brahmagupta's may not be the first name that comes to mind as a geometric algebraist, this interplay – between circle geometry and

(rather simple) binary quadratic forms in modern terminology – is in fact the connecting link between his cyclic geometry and his number theory as we shall see later.

A final remark concerns the uses of the diagonal theorem. The theorem in its direct form is employed in the *Śulbasūtra*, for instance in doubling the area of a square or, more generally, in scaling it up n-fold. But a large number of constructions – e.g., rational right triangles by one of the cord methods – are actually based on its converse in the form: if the sides a, b, c of a triangle satisfy $c^2 = a^2 + b^2$, then the angle opposite c is the right angle. Curiously, the converse theorem is not stated in any of the *Śulbasūtra* and no statement, let alone a proof, is given (as far as I know) in any of the later texts either.

2.3 Rectilinear Figures and their Transformations

The basic geometric theme of the *Śulbasūtra* is the transformation of figures into one another with a given relationship between the areas of the initial and final figures. More explicitly and with one significant exception (taken up in the next two sections), they are transformations of polygons to polygons of the same or a different form, having a degree of symmetry (squares and rectangles, isosceles trapezia, parallelograms and rhombuses) corresponding to the different shapes of the canonical altars. The methods of construction often mix the purely geometric (compass and straightedge) with some amount of mensuration. Many procedures are of the cut-and-move or area-matching variety but quite a few invoke the theorem of the diagonal and its converse. The examples I describe, some of them very briefly, are meant to be illustrative rather than exhaustive. Along the way, I will try to distil out of this diverse collection of exercises the basic geometric truths the authors of the *Śulbasūtra* may have discovered, i.e., apart from the diagonal theorem.

A large number of the constructions are common to all the *Śulbasūtra*, as is natural considering their purpose in rituals. There are variations in their ordering from text to text, in the formulation of the instructions, in the methods employed and in the relative importance given to the roles of geometric or mensurational aspects. But there is always an underlying unity of theme and method; in particular it is only rarely that we can glimpse, perhaps, some significant change in mathematical thought over the four centuries or so the texts are generally supposed to cover.

Before taking up area-scaling transformations, it seems natural to first deal with the making of the most basic of the figures (apart from the circle itself), namely squares and rectangles of given linear dimensions. Triangles apparently had less of ritual importance at that time – there is one isosceles triangular altar, the *praugaciti*, with a base of 2 units and symmetric side of $\sqrt{5}$ units, made by joining the midpoint of one side of a square to the two opposite vertices – and and they occur generally as a result of cutting up other figures or as an intermediate step. This natural order is respected by Baudhayana but

not always by the other authors. Ritual symbolism also required the axis of symmetry to be along the east-west direction. Once the east-west line is determined, the basic geometric step required to be mastered is the drawing of the perpendicular to a given line through a given point on it. One obvious way is to use cord-and-peg methods based on the converse theorem of the diagonal, as described in the previous section in connection with scaling. But more interesting are methods not involving measurement (except for the overall size) or calculation, but depending on geometric principles other than the diagonal theorem: the altitude of an isosceles triangle to the base (the unequal side) divides it equally, or: the common chord of two equal intersecting circles is the perpendicular bisector of the line joining their centres and is itself divided equally. The two formulations correspond to two different operational methods, even though the first is only a particular expression of the more general second formulation. In the first method, a long enough string with a knot at the midpoint is fixed at its ends and stretched transversally by pulling on the knot and, in the second, equal circular arcs are drawn with a cord-compass with two equidistant points as centres, as we do today. The latter, more intrinsically geometric, method is the first step in Baudhayana's construction of a square and what we can learn from it is best dealt with after going through that construction.

Setting aside the paraphernalia of bamboo rods and strings, Baudhayana's first method for constructing a square of side $2a$ proceeds as follows (Figure 2.7). Let E and W be two points $2a$ apart on the east-west line and let O be the midpoint of EW. (In the text, O is found by measurement, i.e., by folding the cord EW in two.) First, draw a circle with centre O and radius a. Next, draw two circles of radius $2a$ with E and W as centres and let them intersect at N' and S' (with N' to the left of EW in the diagram). The line $N'S'$ which passes through O intersects the first circle at N and S. Now draw circles centred at E, N, W and S, of radius a. All of them pass through O and intersect pairwise at A, B, C and D. Then $ABCD$ is the required square.

Before we even think of the geometry behind the construction, what leaps to the eye is the pleasingly symmetric geometric design the various steps have together created. I will have a great deal more to say about the pattern-generating potential of the diagram of Figure 2.7, and about where the method might have originated from, in the next chapter.

Given only a pair of compasses and a fixed line for orientation, this is the simplest possible way of making a square of a given side and (almost) the most economical – almost, because the initial step of finding O by measurement is avoidable as it is the intersection of $N'S'$ and EW. It is particularly to be noted that the diagonal theorem has no role in the construction. Instead, the geometric principle it is based on is the mutual orthogonality of the common chord and the line of centres of intersecting circles to ensure that NS is orthogonal to EW. There is a second principle involved: in the quadrilateral $AEON$, all the sides are equal by construction and the sides OE and ON are orthogonal; hence it is a square. The second principle can be stated in other equivalent ways, but this is intuitively the most direct.

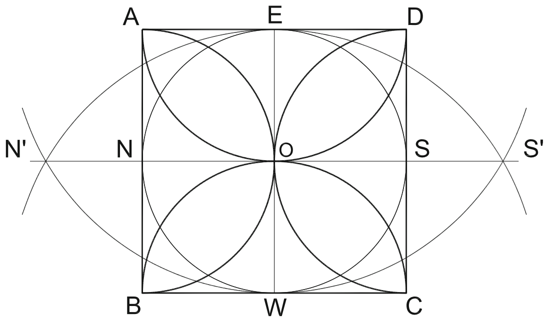

Figure 2.7: Baudhayana's cord-compass construction of the square

Some other remarks are in order. Firstly, if the requirement of cardinal alignment is dispensed with, $ENWS$ is already a square, but of side $\sqrt{2}a$. Since the dimension had a ritual significance, such a deviation would not have been acceptable. At first sight, it may appear that the shortcoming can be compensated for by a technique much discussed and used in the *Śulbasūtra*, that of doubling the area of the square by making a square on the diagonal. That will not do as area-doubling is itself based on constructing a square on a given line as described above. It will also entail the use of the theorem of the diagonal. That may be an important detail as it reinforces the idea that the construction dates back to the Indus Valley; as I will argue in the next chapter, there is no evidence that the Harappans knew the diagonal theorem.

The second remark is that the radius of the big circles centred at E and W and intersecting at N' and S' could have been chosen to be any value greater than a without affecting the final outcome. But it would have changed the (Harappan) geometric design generated by its repetition, making it less symmetric and without the same striking visual appeal.

Several other methods of drawing squares are described in all four of the texts, based on one or another way of drawing perpendiculars, generally involving the diagonal theorem and diagonal triples and/or mensuration; there is even one in which a cord of length $\sqrt{2}$ is to be prepared, presumably by measuring the diagonal of the unit square. Also, once perpendiculars are drawn, such methods can easily be adapted to make a rectangle of any desired sides, which in turn can be turned into isosceles trapezia by an elementary construction.

Of all the geometric transformations treated in the *Śulbasūtra* the majority are those preserving areas. Many of them are straightforward applications of area-matching and hold no new geometric interest. The example discussed next – converting rectangles into squares and conversely – is one which has some points of novelty and in which the construction is not completely obvious.

Let $ABCD$ be a rectangle of sides a and b with $a > b$ (Figure 2.8). Mark M and N on the long sides such that $AMND$ is a square (of side b) and the midpoints P and Q of MB and NC. Cut out the rectangle $PBCQ$ (of length b and width $(a-b)/2$) and move it to the position $ARSM$. Then the figure

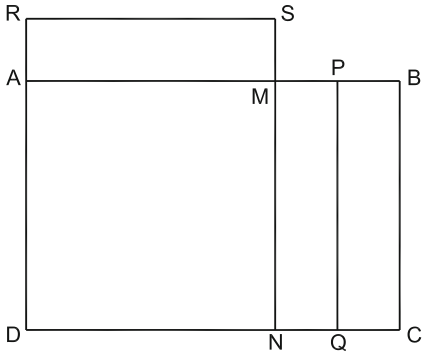

Figure 2.8: Rectangle into square

$DRSMPQ(D)$ has the same area as the original rectangle. It is not a square but the 'difference' of the square on DQ (of side $b + (a - b)/2 = (a + b)/2$) and the square on MP (of side $(a - b)/2$), corresponding, after multiplication by 4, to the identity $(a + b)^2 - (a - b)^2 = 4ab$. There is a purely geometric procedure for producing squares of areas which are sums and differences of areas of squares (described below), whose application leads to the required square.

The converse problem of converting a square into a rectangle with one side given is interesting because it is posed (by Apastamba) but, effectively, not solved. The solution came very much later from Sundararaja, one of the commentators of Apastamba and a contemporary of Nilakantha. Suppose the given side of the rectangle is the long side $b(> a$, the side of the square $ABCD)$. Extend DA and CB to M and N such that the length of DM and CN is b and let CM intersect AB at P (Figure 2.9). The line through P perpendicular to AB then gives rise to two rectangles $APQD$ and $BPRN$ (see the figure). Since the line CM is a common diagonal of the three rectangles $MNCD$, $MRPA$ and $PBCQ$, area matching shows that the rectangles $APQD$ and $BPRN$ have the same area. Hence the rectangle $RNCQ$ and the square $ABCD$ have the same area. Equivalently, if the length of PB and QC is denoted by c, the algebraic condition expressing this equality is $c(b - a) = a(a - c)$ or $a^2 = bc$. A minor change in the procedure takes care of the case when it is the short side of the rectangle that is given.

It is a disappointment to find that, after having posed the problem, Apastamba leaves it without a solution – or gives such an ambiguous prescription as to amount to the same thing (see [SA]) – since Sundararaja's elegant area-matching is so clearly in the *Śulbasūtra* tradition. But Katyayana, later than Apastamba, does describe a method of turning the unit square into a rectangle with the long side equal to $\sqrt{2}$, which is a minor modification of the suggested proofs of the diagonal theorem for the square (Figure 2.10 needs no further explanation). Indeed one may put forward Katyayana's construction as yet another piece of circumstantial evidence that the *Śulbasūtra* authors did have some such "self-evident" justification for the theorem.

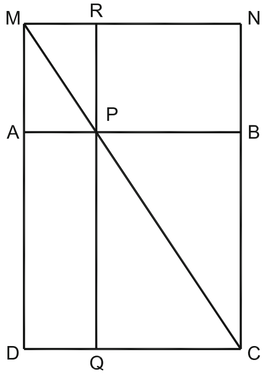

Figure 2.9: Square into rectangle

The *Śulbasūtra* also treat (approximately) area-preserving transformations which are inherently not linear, namely between circles and squares These quite remarkable constructions hold enough points of geometrical and arithmetical interest to deserve separate sections, after we discuss other linear problems such as 'adding' and 'subtracting' squares (the reference is to their areas) and scaling them up and down.

The most elementary procedure of this type is the addition of two equal squares, the doubling of a square. It occurs freely in all the *Śulbasūtra* and always in the same, most obvious, form: construct the square on the diagonal; in fact, Apastamba states the theorem of the diagonal and the prescription for doubling in the same *sūtra* (and reiterates the prescription later). There is nothing to add except to note the nomenclature: the diagonal of the square is called *dvikaraṇī*, 'two-maker' or 'double-maker'. Adding two arbitrary squares of sides a and b is, just as obviously, equivalent to the general diagonal theorem: make a rectangle of sides a and b and draw the square on the diagonal. In particular, to triple the unit square draw a rectangle of unit width on the diagonal of the unit square as the base and then draw the square on the diagonal of the rectangle; the diagonal is called *trikaraṇī*, 'three-maker'. More generally, if a_n is the length of the diagonal of the rectangle with a as the short side and a_{n-1} as the long side (with $a_1 = \sqrt{2}$), $a_n^2 = (n+1)a^2$. This construction has a special point of interest: it is the first application of a recursive principle in geometry.

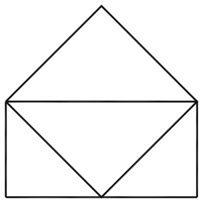

Figure 2.10: Square into rectangle and the diagonal theorem for the square

Subtracting a square from a larger one is also a direct application of the diagonal theorem and need not be detailed.

The idea of drawing repeated diagonals for multiplying a square is present already in the earlier *Śulbasūtra*. In his approach to the same problem, Katyayana, 300 - 400 (?) years later, introduces a genuine innovation. The *sūtra* is obscurely written. Under a favourable reading, the instruction is simple and general: make an isosceles triangle of base $(n-1)$ and (symmetric) side $(n+1)/2$; its altitude is the side of the required square for, by the diagonal theorem, the square of the altitude is $((n+1)^2-(n-1)^2)/4=n$. Katyayana here does away with the mechanical repetition of a well-understood procedure and replaces it with a one-step construction valid for any n, probably the first time that such a move towards the abstract is made in India. Parts of the *sūtra* may be obscure – it says nowhere that the triangle is isosceles – but there is no doubt about the meaning: "as many squares as you wish to combine into one" (Bibhuti Bhushan Datta's translation, quoted in [SA]; the translation in [SuSu-SB] is not materially different). In any case, the ambiguity is not fatal to the conclusion, as is evident from the fact that the altitude is taken to divide the base equally; the correctness of the construction is the guarantee of the correctness of the reading. It is also clear that by the time of Katyayana, algebraic identities of the type employed here had outgrown their geometric roots and were understood to be valid for any numbers which could be multiplied and added, arbitrary positive integers in this example.

It is reasonable to consider this particular construction of Katyayana (whom most commentators date after Panini) as a sign of the evolving mathematical culture, even though the substance of his *Śulbasūtra* is still much the same; rituals had still to be performed and they still had to conform to the ancient canon. There is another instance of a similarly shifting point of view and this time it is in the way the diagonal theorem itself is formulated, not by Katyayana but by Manava (who is later than Baudhayana and Apastamba, but how much later?). It starts with "Multiply the length by the length and the breadth by the breadth ...." and concludes that the length of the diagonal is the square root of their sum (see [SuSu-SB]). The theorem is no longer primarily about areas but about numbers and arithmetical (algebraic) operations (multiplication, square roots) on them. As we shall see in the next section, Baudhayana and Apastamba already had a good arithmetical understanding of $\sqrt{2}$, but the shift away from areas to lengths as magnitudes, to be dealt with as any other numbers, marks another step forward in the early evolution of mathematical thought.

2.4 Circle from Square: The Direct Construction

A few of the Vedic altars had a circular shape for the base; so it is natural that area-preserving transformations involving circles interested the ritual experts from the beginning. Since the circular perimeters were constructed out

of bricks having the form of a symmetric trapezium with the shorter parallel side (the 'face') facing the centre, the circles were never exact circles, but only polygonal approximations to them, a fact that could not have been missed by people whose first geometric instinct was to draw a circle with a pair of cord-compasses. But before placing the bricks so as to form an approximation to a circle, an actual circle was presumably drawn and the geometric transformations have to do with such ideal altars. The reason for mentioning this obvious point is that historians sometimes tend to forget that the theory of ritual allowed the ideal, when it was unattainable, to be replaced by the best substitute to hand. For example the substitution of the plant from which the intoxicating-hallucinogenic libation *soma* was extracted – whatever it was in a quasi-mythical past – by one without any 'magical' qualities was authorised already in middle Vedic times. The comparison is not so far-fetched; though not deified in the Vedas, the geometry of the altars had a divine significance and had as much symbolic-magical relevance in the context of ritual as did *soma*. In any case, much intellectual effort went into the construction of the circle having the area of a given square and into the converse construction.

No part of the *Śulbasūtra* has engaged the scrutiny of modern commentators as extensively or attracted as much imaginative conjecture as the material on the circle. That is to be expected. In a situation where the textual descriptions are sometimes opaque and where we cannot always tell whether the opacity resulted from sloppy writing or poor understanding, there is bound to be controversy. Apart from the diverse rationales offered for some of the statements, there are disagreements among modern scholars on how some of the *sūtra*s are to be translated and read and even on whether one or two of them make any (literal) sense at all. The account given below is intended to highlight those aspects on which there is a general consensus and to indicate where there is a lack of unanimity and why. Even with such a modest goal, and without direct textual evidence of what geometrical facts the *Śulbasūtra* authors took for granted, we have to extrapolate. It will therefore be helpful to first try and form a picture, even if somewhat speculative, of the state of geometric knowledge of the time concerning the circle.

The *Śulbasūtra* geometers say hardly anything about the connection between the circumference C of a circle and its radius r, areas being the geometric quantity that most mattered to them. (There is one place where Manava says that the cirumference is a numerical multiple of the diameter; Baudhayana himself takes the circumference of a circular pit to be three times its diameter, also exactly once). In any case, since all ancient civilisations seem to have realised that the two are proportional, it would not be rash to impute that knowledge to the early (pre-Manava) Indian geometers as well. An intuitive way to get there is to take the circle as one circumscribing a regular polygon with a large number of sides and then to use the proportionality (scaling) of the sides of similar triangles (see Figure 2.11). We have seen that the scaling property is used in some *Śulbasūtra* constructions and the circular altar itself was made as a regular polygon with as many sides as the number of bricks used for the

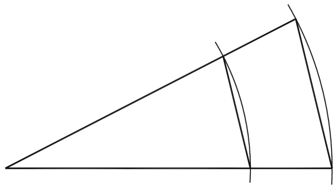

Figure 2.11: Scaling of the circumference with the radius by triangulation

perimeter. The estimation of the area A of the circle in terms of the radius by the polygonal approximation is more complicated and less intuitive; it could have been done in principle from the (known) area of the right triangle but at the cost of having to take many nested square roots.

But, as we have also seen, the Vedic geometers certainly knew that areas of simple rectlinear figures scaled as the square of some linear dimension and could, naturally, have extrapolated that insight to the circle (which is determined by one length, the radius): $A = Kr^2$ for K some number. We shall see in this section that this knowledge is implicit in their approach to area-equivalent squares and circles. Moreover, by drawing two squares, one inscribed in the circle (of diagonal $2r$) and the other enclosing the circle (of side $2r$) (Figure 2.12 below) they would have known that A is between $2r^2$ and $4r^2$. There is no evidence that they knew that K is also $C/2r$.

It would also seem absolutely safe to suppose that their favourite method of matching areas could not have been made to work in the case of the circle, because of the curvature. Interestingly, when an exact formula for the area was arrived at later in India, it was expressed as $A = (1/2)Cr$ (in *Āryabhaṭīya*, probably known from a little earlier), the area of the rectangle having half the circumference and the radius as the long and short sides. (Area-equivalent circles and rectangles do not figure in the *Śulbasūtra*). It was almost certainly derived by a cut-and-move method but it also needed a new idea for dealing with curvature, that of dividing the circumference into a very large number of equal pieces and ignoring the effect of the curvature in the very small arc segments thus formed; in other words, by approximating the circle by a regular polygon of n sides and taking the limit of large n (see Chapter 7.5). Needless to say, the Vedic geometers are unlikely to have thought along such essentially Aryabhatan lines.

So, the question is: how accurate were the *Śulbasūtra* ideas of the value of K and how were they arrive at? The question is important because a good guess at the value of K is essential for even contemplating the problem of transforming a square into a circle of (approximately, even in principle, as we know now) the same area; unless one was satisfied by rough visual comparisons, there would have been no benchmark without it, no way of quantifying the goodness of the approximation. To put it another (computational) way, the problem involves deciding on the values of two parameters: the constant K, fixed but supposed unknown within its bounds and, for a given K, the equivalent radius. If the

value of K were known, the radius is determined as a square root and there is a technique of recursive approximations for dealing with square roots in the *Śulbasūtra* themselves (which we shall meet soon). If K is numerically unknown, geometry cannot determine the equivalent radius.

All four of the *Śulbasūtra* have prescriptions for the construction of a circle having the same area as a given square (which, for ease of reference, I will call the direct construction). Baudhayana's method is later restated by Apastamba and Katyayana, in substantially the same terms; Manava has an entirely different approach to the problem. None, of course, provides anything remotely like a rationale. Manava is slightly more accurate in the final result but the procedure involves two arbitrary parameters, fixed judiciously (assuming that he has been read correctly), and appears artificial and forced; it is not organic to the circle and is not a modification of Baudhayana's method: even the first step is different. We shall only be concerned with the Baudhayana version because his converse construction of the square equivalent to a circle is best regarded as its inversion and also because its simplicity lends itself more readily to conjectural theorising. In the absence of any useful insight, ancient or modern, into the thinking behind the method, any reasonably credible guess should be worth a look.

There have been many readings of the passage prescribing the construction and they all agree on its interpretation. It can be paraphrased as follows. Let $ABCD$ be the given square, of side 2 units, and O its centre, presumably known from the way it was constructed. Draw the circumscribing circle of the square. Let P be the midpoint of AB and let OP intersect the circle at Q. Mark the point R on OP such that PR is $\frac{1}{3}PQ$. Then the required circle is the one having O as centre and passing through R (Figure 2.12). Since the radii of the inscribed and circumscribing circles are 1 and $\sqrt{2}$, it is immediately obvious that the construction respects the upper and lower bounds on K: $2 < K < 4$, because the correct circle (no matter that we cannot construct it exactly) will intersect PQ somewhere between P and Q. That is all that can be assumed known with certainty from the geometry.

It is a striking fact that while the above construction is strictly geometric (apart from the measurement of the one-third part of PQ, that is), the con-

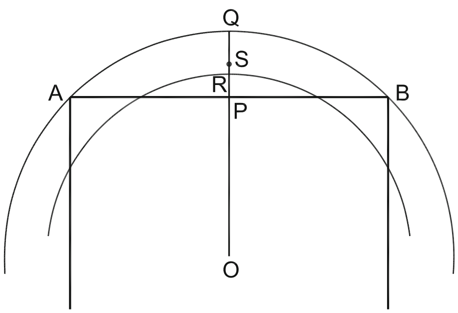

Figure 2.12: Baudhayana's construction of the circle equivalent to a square

verse construction of the square equivalent to a given circle, to be taken up next, is expressed primarily in arithmetical terms. The needed measurements are all given in numerical formulae involving negatives and fractions with fairly large denominators (and, behind the scenes, square roots). That the *Śulbasūtra* geometers were adept at routine operations with decimal numbers is already evident from the considerations of the previous section. That should not be too surprising: at least four centuries separated them from the *Ṛgveda* whose number names, as we shall see in later chapters, not only show a mastery of the principles governing the formation of decimal numbers but also an understanding of basic addition and multiplication.

It would be somewhat unnatural to imagine that Baudhayana thought of the two mutually inverse problems from such different mathematical points of view, one geometric and the other numerical, as though the problems were themselves fundamentally different in nature. If we accept that, though the description of the construction of the circle is geometric – what else could it be given that there is only one number, 1/3, that figures in it? – the reasoning that led to it could have been computational, we can begin to explore non-geometric paths that might have led to the final result. I outline below one such line of reasoning.

The technique, very Indian and of great antiquity, is a general way of dealing with deviations from linearity by a method of recursive approximations. Whenever a quantity is to be determined from three others which are not in relative proportion – i.e., when the rule of three or *trairāśikam* is not applicable – the method consists, in its most elementary form, of making a first guess at the answer and then refining it by a sequence of linearly determined corrections. The first guess is often chosen to be the result in the linear approximation, i.e., from the (invalid) application of the rule of three, but that is not compulsary. The refining can go on indefinitely; the more astute the starting guess, the fewer are the steps needed for a given level of accuracy in the answer. A perfect candidate for the use of the method is the calculation of the square root of a non-square number and the resulting optimal formula after the first two steps of refining is in the Bakhshali manuscript. Leaving an account of how it works in general for later, let us straightaway turn to Baudhayana's problem.

As noted earlier, there are two facts that Baudhayana undoubtedly knew: i) the point R on the line OP through which the true equivalent circle will pass lies between P and Q; ii) the areas of the circles through P and Q are in the ratio 1:2. If the rule of three held, which of course it does not for radii and areas, R would be the midpoint S of PQ; so S will be the first guess for R, with $OS = (1 + \sqrt{2})/2$. Knowing also that the area goes like (radius)^2, faster than linearly, Baudhayana would conclude that the circle through S overestimates the desired area of 4. The first step in refining this crude estimate is then to suppose that OS has to be corrected by a small negative quantity $-\delta(\sqrt{2} - 1)$, $\delta > 0$, where I have expressed the correction as a multiple of $PQ = \sqrt{2} - 1$ since that is how Baudhayana specifies the point R. The method of refining now requires that the corrected radius must satisfy the condition that the area

determined by it is the area of the square,

$$K\left(\frac{1+\sqrt{2}}{2}-\delta(\sqrt{2}-1)\right)^2 = 4.$$

Assume now that δ is small enough to ignore the term proportional to its square (linearisation), leading to

$$\delta = \frac{3}{4}+\frac{1}{\sqrt{2}}-\frac{4}{K}.$$

This is an elementary but interesting equation because it illustrates the point that without a good idea of the value of K, we cannot hope to get a good approximation to the true position of Baudhayana's point R (and without any idea of the value of K, the equation is useless). For the two limiting values $K = 2$ and 4, δ turns out, in decimal fractions, to be approximately 0.45 (very close to the incircle) and -0.55 (outside the circumcircle) respectively. Clearly, one has to do better.

In the spirit of computing corrections to *trairāśikam*, let us now make a first guess of the value of K as the mean of the two geometrically obvious limiting values, namely, $K = 3$. Baudhayana was certainly aware that the correct value is close to this; in *sūtra* 4.15 he says, while specifying the dimensions of circular pits, that the circumference of a circle of diameter 1 is 3. That leads (reverting to decimal fractions and using $\sqrt{2} = 1.414\cdots$) to the value $\delta = 0.194\cdots$. This is the approximate distance, as given by the refining process, of R from S. The corresponding length of PR is then $0.5-0.194\cdots = 0.306\cdots$ (in units of PQ), to be compared with Baudhayana's simple fraction 1/3. It is clear that slight adjustments in the value of K will produce an answer as close as we wish to 1/3; obviously, there is not much point in doing it. We can also repeat the process with $(1+\sqrt{2})/2-\delta$ as the new input but the only point of doing that will be to convince ourselves that the new correction will be much smaller than δ.

It must be stressed, again, that Baudhayana was not trying to compute a value for π (that is the reason I have chosen not to denote the approximately known constant K by π), but only to characterise the equivalence of squares and circles with the means he had at his disposal, including a reasonable idea of the value of K. It must also be kept in mind that what I have described is no more than an attempt at second-guessing Baudhayana; like all such attempts, there is always the unavoidable risk of it being off the mark. But the fact that the result comes close to Baudhayana's 1/3 is not accidental. The closeness derives from two sources. Firstly, Baudhayana's choice of K, which is a matter of prior knowledge, cannot have been far from 3 and, secondly, once that choice is made, the logic of the process of refining guarantees ultimate success in the sense that, if the value of δ obtained by linearising as above is not good enough for our taste, the process can be repeated to obtain a new correction – presumably, a radius of $1+(\sqrt{2}-1)/3$ was thought good enough. These points will play an even more significant role in the converse construction.

2.5 The Inverse Formula: Square from Circle

All the texts give a value for the side a of a square having area equal to that of a circle of diameter d, $a = (13/15)d$. Numerically this does not lead to a very good approximation to the correct equivalent square. Its interest lies elsewhere: the statement is no longer about a step-by-step construction at all but about the numerical ratio between appropriate linear dimensions of area-equivalent circles and squares; the equation $a = (13/15)d$ can just as readily be used to make a circle from a given square (the direct problem) as the other way round. Stated differently, we are not dealing with a problem of cord geometry any more but with an attempt to reduce the geometry to formulae and to computations with them.

Already we have here a hint that the direct construction was seen as the more fundamental problem, to be solved first, geometrically (or, maybe, by a combination of geometry and computation). Such an interpretation is consistent with the fact that the building of no ritual structure required the transformation of the circle into any other figure. Once the direct construction was accomplished, there were options to tackle the inverse problem. The geometrical way is to start from first principles as in the direct problem, say by bounding the circle by two squares of which it is the circumcircle and the incircle and, perhaps, computing corrections to the result of applying the rule of three. That route was not followed. Another very simple option (which requires no knowledge of the direct construction and hardly any geometry even) is to assume a value of K and match areas by means of formulae: if $a = \kappa d$ for a constant κ, we have $a^2 = K(d/2)^2 = (\kappa d)^2$. For the value $\kappa = 13/15$, the only one found in all four texts, this gives $K = (2\kappa)^2 = (26/15)^2 = 2.99\cdots$, as close to 3 as makes no difference (given the distaste of Indian mathematicians for decimal fractions). In fact Sen and Bag ([SuSu-SB]) suggest in their commentary that $13/15$ is the value of κ that results from $K = 3$ (so that $a = (\sqrt{3}/2)d$) and estimating $\sqrt{3}$ by a process of refining very much like the one discussed below for $\sqrt{2}$ and in the direct construction given above for Baudhayana's equivalent radius. Perhaps this is evidence that 3 was the commonly accepted value of K at the time of Baudhayana.

But Baudhayana apparently had second thoughts. He (alone) has another and much more elaborate numerical expression for the value of κ:

$$\kappa = 1 - \frac{1}{8} + \frac{1}{8 \times 29} - \frac{1}{6 \times 8 \times 29} + \frac{1}{8 \times 6 \times 8 \times 29}.$$

This is quite an amazing combination of fractions; not surprisingly, no one (to my knowledge) has given a geometrical interpretation of it. A reasonable hypothesis is that Baudhayana must have exercised yet another option: start from his own result (the direct construction) for the radius of the equivalent circle and work backwards. There is some *ab initio* support for such a likelihood in that he also gives a numerical expression for $\sqrt{2}$ in this connection (as another

remarkable combination of fractions):

$$\sqrt{2} = 1 + \frac{1}{3} + \frac{1}{3 \times 4} - \frac{1}{3 \times 4 \times 34}.$$

As we have seen, $\sqrt{2}$ figures prominently in the direct construction – in fact it will figure in any circular approximation to the square that begins by invoking the inscribed or the circumscribing circle – even though fundamentally there is no reason for it to appear in what is in principle a calculation involving π.

It may well be that the purpose of the inverse construction is as a means of checking the area of the circle constructed by the direct transformation. It serves no ritual purpose and Baudhayana's concern that it be an accurate inversion of the direct construction is consistent with such an interpretation. That will be interesting if true, because it will mean that no direct way of accurately determining the area of the circle (the value of K) was known. There is at least one other instance of an inverse construction being resorted to for a similar reason, Apastamba's conversion of a trapezium into a rectangle in order to determine the area of the former (which is the form the great arena, *mahāvedi*, takes).

Now, since κ is defined as a/d, it is the reciprocal of the diameter d_1 of the circle equivalent to the unit square. From the direct construction, $d_1 = 1 + (\sqrt{2} - 1)/3 = (2 + \sqrt{2})/3$; hence, if the hypothesis that Baudhayana just inverted the direct construction is correct, $\kappa = 3/(2 + \sqrt{2})$. This more intricate inversion will still be a purely arithmetical exercise and the challenge is, using Baudhayana's rational approximation for $\sqrt{2}$, to organise the computation involved in such a way as to enable a comparison with the expression for κ in the particular fractional form cited above. From Thibaut's pioneering work in the 1870s onwards, there have been several attempts to meet the challenge, the latest being in a recent article by Dani[2]. In making explicit the connection between Baudhayana's radius of the equivalent circle and the expression above for κ, *via* $\sqrt{2}$, all of them recognise that the inversion is not exact and therefore need to resort to some clever approximation or the other.

As an example and in summary, Dani's suggested rationale goes as follows. Start with $d_1 = (2+\sqrt{2})/3$. On substituting the rational approximation for $\sqrt{2}$ given above, this becomes $d_1 = 1393/1224$; so $\kappa = 1224/1393$, which however cannot be written as a sum of fractions with smaller denominators as 1393 is a prime. Dani suggests that a fraction $7/8$ was removed from this and that, in the remainder which is $41/(8\times1393)$, 1393 was adjusted to 1392. Now 1392 factorises as $6 \times 8 \times 29$ and in fact $41/1392 = 1/29 - 1/(6 \times 29) + 1/(6 \times 8 \times 29)$.

This reworking makes one thing absolutely clear: the only possible sources of significant error are Baudhayana's equivalent radius and the approximation

[2] S. G. Dani, "Geometry in the Śulvasūtras", in [SHIM]. This reference also has an interpretation of Manava's unorthodox direct construction as well as a critical appraisal of the various attempts to derive the assertions regarding the square-circle connection in the *Śulbasūtra*.

for $\sqrt{2}$. The numerical formula for $\sqrt{2}$ is in fact extremely accurate (correct to five significant figures); therefore the value of κ is as good or as bad as the direct construction is. In particular, K does not play any direct role in the logic – its influence is through the direct construction – and we cannot learn anything new from the inverse formula about how accurately it was known in Baudhayana's time. We have already seen that the direct construction is perfectly consistent with a value close to 3 and there the matter will have to rest.

A second remark is that all the reconstructions presuppose a proficiency in the arithmetic of fractions at least of the level invoked above. If these reconstructions do reflect the process by which the *Śulbasūtra* authors worked out the necessary arithmetic, their computational skills must have been of a high order: adding and subtracting many fractions by finding the common denominator, factorising fairly large numbers and recognising when they are prime (involving division by two-digit numbers), etc. This example alone should be enough to dispel the notion that they were constructive geometers who were, by implication, not so skilled in numbers and arithmetic. On the contrary, it would seem that they did not shy away from the 'metric' in geometry and had a pretty good feeling for when and where a particular method, geometric or arithmetical or both in combination, was the most suitable.

That having been said, mysteries remain. There is a very simple and direct way of incorporating the Baudhayana approximation for $\sqrt{2}$ in his expression for the diameter leading, with no further adjustment, to a formula for κ as a sum of fractions. Moreover, it agrees with the cited expression in the first two terms and is far easier to use (no really large denominators). In the formula $\kappa = 3/(2+\sqrt{2})$, just rationalise the denominator, giving $\kappa = (3/2)(2-\sqrt{2})$, and substitute for $\sqrt{2}$:

$$\kappa = \frac{3}{2}\left(2 - 1 - \frac{1}{3} - \frac{1}{3\times 4} + \frac{1}{3\times 4\times 34}\right) = 1 - \frac{1}{8} + \frac{1}{8\times 34}.$$

Why is this more accurate value of κ, less cumbersome and easier to derive, not given by Baudhayana? It cannot be because he could not manage the arithmetic or did not know about rationalisation which after all is about cancelling common factors in the numerator and the denominator. We are not really in a position to say that everything is understood.

Finally, there is the question of how $\sqrt{2}$ was worked out. Here again, in true *Śulbasūtra* fashion, the geometric and arithmetical aspects come together as two facets of the same idea. First, here is an arithmetical approach using, once more, the method of recursive refining. Suppose we approximate $\sqrt{2}$ by, reasonably, $1 + 1/3 = (\sqrt{2})_1$ as the first guess and let the error be δ, $(1 + \frac{1}{3} + \delta)^2 = 2$. Neglecting δ^2, we have a linear equation for δ: $\frac{16}{9} + \frac{8}{3}\delta = 2$ which is solved by $\delta = \frac{3}{8}\left(2 - \frac{16}{9}\right) = \frac{1}{3\times 4}$. The output value is $(\sqrt{2})_2 = 1 + \frac{1}{3} + \frac{1}{3\times 4}$; use it as the new input, with δ' as the second correction and solve for δ', again in the linear approximation. The answer, as is easily checked, is $\delta' = -\frac{1}{3\times 4\times 34}$, leading to a value for $(\sqrt{2})_3$ identical to the *Śulbasūtra* expression.

The game can be varied by changing the first guess. If we take $(\sqrt{2})_1$ to be 1, the first two correction terms are $1/2$ and $-1/12$, which add up to $5/12$, which is the same as the sum of $1/3$ and $1/12$. It is clear from the process that if two different first guesses lead, each after a (different) finite number of steps, to the same output, all further corrections will coincide from then on, since the inputs will be the same. Even more fun: start with 2 as the first guess (absurd as it may seem to think of 2 as its own square root); then the first correction is $-1/2$ and that is enough since $1 + 1/2 = 2 - 1/2$.

We have come across enough instances of the interplay between elementary quadratic algebraic identities and plane geometry not to doubt that extracting square roots by successive refining will also have its geometric counterpart: just represent squares and products of numbers by suitably configured squares and rectangles and match areas. But let us first note that, of the two conceptual ingredients of this procedure, the more routine, the approximation $(a + b)^2 = a^2 + 2ab$ (the truly original idea being its iteration) is equivalent to what is generally called Heron's formula for the square root, $\sqrt{a^2 + b} = a + b/2a$. The suggestion that the *Śulbasūtra* $\sqrt{2}$ was obtained by the use of Heron's formula, extended by including the next (binomial) correction term, was made already in the 19th century (by Léon Rodet, described in detail in [SuSu-SB]). As we have seen, the *Śulbasūtra* mathematicians would not have needed to know the next term in a putative binomial expansion of square roots of sums for this purpose, but only how to iterate the process of finding the first term and, of course, how to manage the basic arithmetic of fractions.

The geometrisation of the elementary identity $(a+b)^2 = a^2+2ab+b^2$ occurs in many rectlinear constructions in the *Śulbasūtra*, some of which we have come across in an earlier section. Whenever the area of the square on b can be neglected it is, automatically, the geometric version of Heron's formula as well. More to the point, the iterative square root also has an obvious geometric equivalent which the reader can easily supply, in fact several equivalents depending on the first guess. The ingenious geometric demonstration of the value of the *Śulbasūtra* $\sqrt{2}$ put forward by Datta and explained in detail in [SuSu-SB] (see also Dani, cited above) is essentially this geometrisation for $1{+}1/3$ as the first guess: the starting points of both Rodet and Datta are obviously chosen to agree with the first two terms. In any case, the arithmetical method is elementary, easy to implement and well within the capacity of the experts who first put together the *Śulbasūtra*, if we go by the skills they have demonstrated elsewhere in the texts, especially in the derivation of the inverse formula.

The refining process can be continued indefinitely for any desired level of accuracy but the answer after any finite number of steps is never exact. There is a strong indication that this was understood. Baudhayana concludes the *sūtra* for the value of $\sqrt{2}$ with the qualification *saviśeṣaḥ* and both Apastamba and Katyayana express similar qualifications, all based on the word *viśeṣa*. Now *śeṣa* means 'that which remains' and the most direct interpretation of the qualifying phrases is that between $\sqrt{2}$ and the fractional sum given as its value, there is something left over, a remainder. In the light of what we have seen above, it

is tempting to suggest that the reference is to the fact that recursive refining generally leaves a remainder; in other words, the qualifying phrase simply means that the expression is approximate. It certainly does not mean that Baudhayana knew that $\sqrt{2}$ (or π for that matter) is irrational; it is perfectly easy to set up, for instance, an infinite geometric series for a rational number by iterated refining as Nilakantha did, much later, as an illustration of the method. Another similar qualifier (that occurs in the *sūtra* immediately after Apastamba's *sūtra* on the direct construction) is also interesting, but linguistically, in the way it illustrates the occasional (rare) difficulty one can encounter in undoing nominal composition. The word is *sānityā* and the ambiguity arises because it can result from the *sandhi* of both the short *sa* and the long *sā* with *anityā*; in the latter reading, it means that the construction is approximate. (It is difficult to believe that Apastamba might have thought it to be exact). Grammatically, it is a rare example of the nominal composition map being not one-one, without a unique inverse. So much for the infallibility of grammar.

As for the idea itself of recursive refining, recursive structures are a constant presence not only in Indian mathematics[3], but in every aspect of Indian thought and creativity, from the structure of Vedic rituals, to language and grammar – we have already glimpsed examples of structured repetetive sound patterns in our brief look at *padapāṭha* and *prātiśākhya* in Chapter 1 – to music, to myths and literature and so on, throughout India's recorded history. It should therefore occasion no surprise that the first conscious employment of the idea in mathematics could be almost as old as the Vedas.

To conclude this chapter, it is evident from our reading of their mathematical content that the *Śulbasūtra* as a whole present a coherent view of the geometry of rectlinear figures and circles. The creators of this body of knowledge were not mere builders of altars and manipulators of cords; they were enquirers concerned with the principles, fundamentally geometric, that lay behind the practice of sacred architecture. And they brought genuine mathematical insights to that enquiry, the greatest of these insights being, of course, the theorem of the diagonal. But when the problems they faced (or posed themselves, as in the conversion of a circle into the equivalent square) demanded methods not reducible to cord – i.e., compass and straightedge – geometry, they were able, and just as willing, to turn to nongeometric techniques, both numerical and algebraic, to achieve their ends.[4]

One reason the arithmetical content of the *Śulbasūtra* is underplayed, or remains somewhat implicit, may have to do with their prime function as manuals of instruction rather than as expositions of mathematics. The seers could not possibly avoid telling the artisans on the ground how exactly to draw a

---

[3] P. P. Divakaran, "Notes on *Yuktibhāṣā*: Recursive Methods in Indian Mathematics", in [SHIM].

[4] This openness remained a pervasive feature of Indian geometry till the very end, as we shall see in the rest of the book. Statements such as "Geometric interpretations and demonstrations of algebraic techniques are . . . [not commonplace] in Sanskrit [texts]" ([Pl]) are not in line with what is actually in the texts.

floor plan of given shape and dimensions but there would not have been much point in telling them what calculations went into a particular numerical value for, say, $\sqrt{2}$. It makes sense for the builder of the altar to be told "Take the 34th part of the cord" etc., but folding a cord does not have to be the way the number was actually computed. Perhaps the time has come for us to think of the *Śulbasūtra* as part of the general corpus of Vedic mathematics, not just as Vedic geometry. The genesis of the arithmetical component of that more inclusive characterisation will occupy us in Chapters 4 and 5.

3

Antecedents? Mathematics in the Indus Valley

3.1 Generalities

It will be very unreasonable to suppose that the body of mathematical knowledge reflected in the *Śulbasūtra* arose fully formed in the minds of Baudhayana and his fellow-codifiers of ritual doctrine some time in the 8th or 9th century BCE. Rather, as in the appearance of the Brahmi script in a finished form in the 3rd century BCE, the codification was probably preceded by a long period of maturation of ideas and methods, following perhaps a decisive initial conceptual advance (or advances). As far as the mathematics of enumeration (and basic operations with numbers) is concerned, there is very strong evidence (Chapter 5) that decimal numbers had acquired their definitive nominal form – the one which is still current – by the time the *Ṛgveda* was compiled; the decisive advance must have taken place well before then. In regard to geometry, there is no comparable early source from which we can extract, even if painfully, any kind of precise information about what the early Vedic people knew, and when. There are sporadic mentions of some ritual structures in the two recensions of the *Yajurveda* but no useful geometrical insight can be gleaned from them.

But geometry was all-pervasive in the Indus Valley civilisation. Almost every aspect of the life of the Indus people, from the large-scale planning of the cities to the tiniest functional or ornamental object seems to have been governed by a feeling for order and form and geometric symmetry. One would have thought that archaeologists who have spent close to a century studying its material remains and bringing to light the many regularities of form and measurement they display would have looked to see if there was evidence of their continuation into the culture that succeeded it. Such has not been the case. There has been no shortage of speculative attempts to trace linkages with the yet-to-come Vedic culture in intangible ('ideological') matters like belief

P. P. Divakaran, *The Mathematics of India*, Sources and Studies in the History of Mathematics and Physical Sciences, https://doi.org/10.1007/978-981-13-1774-3_3

systems and cult practices and even language, but the less speculative and more practicable task of looking for continuities in the geometric cultures – actual concrete objects on the one side and accurate descriptions on the other – remains work for the future.

Much the same can be said about the work of the Vedic specialists. From the beginning (in the second half of the 19th century), scholars of the *Śulbasūtra* have speculated freely about the possible roots of Vedic geometric knowledge in earlier but geographically distant civilisations – that means really the Babylonian – or about the mutual influences that might have been at work between Vedic and early Greek geometry, with little tangible evidence in either case. But, of the possibility that the sacred geometry of the *Śulbasūtra* might be descended from the practical geometry of the Indus people – neighbours in space and time – there has only been two mentions (to my knowledge) of a general nature, one in a passage in Sarasvati Amma's book ([SA], Introduction) and the other in Seidenberg's attempt to trace the common roots of *Śulbasūtra* and Greek geometry (see below). And, of a focussed effort to either substantiate or reject such a connection, there has hardly been a word in the literature. Happily, there are signs that things are beginning to change. In a very recent summary of what is known of the mensurational and quantitative aspects of the Harappan world, Kenoyer[1] suggests the possibility (among others) that the cardinal orientation of Indus Valley streets and buildings may have been determined by the solar shadow method, the same (though Kenoyer does not say this) as the one later described in the *Śulbasūtra* (see Chapter 2). More substantively, Sinha[2] and his collaborators have initiated the analysis, in the first work of its kind, of the decorative patterns on seals, potsherds, etc. from the point of view of their symmetries and proposed a way of understanding how they may have been generated as planar lattices from a unit cell.

In the absence of textual material to fall back upon, to think of these observations as anything more than the first strokes in drawing a portrait of the mathematical culture of the Indus people will be overoptimistic. For an appreciation of the challenges such an undertaking will have to meet, just imagine that none of the *Śulbasūtra* texts that we possess now had survived into modern times – as seems to have happened with quite a few of them in any case – and that all that we can know about Vedic geometry has to be deduced from scattered ruins of brick altars. It is true that the Indus civilisation has left us impressive ruins, and a great deal more of them than the Vedic period did, but

[1] J. Mark Kenoyer, "Measuring the Harappan world", in *The Archaelogy of Measurement*, Eds. I. Morley and C. Renfrew, Cambridge University Press (2010). Already in the 1920s, after the first systematic excavations at Harappa and Mohenjo Daro, it was recognised that numerical and geometric regularity was a hallmark of the Indus culture, more than of other contemporary cultures. The results of more recent work at other sites have only reinforced that conclusion.

[2] S. Sinha, N. Yadav and M. Vahia, "In Square Circle: Geometric Knowledge of the Indus Civilization", in *Math Unlimited: Essays in Mathematics* (Eds. R. Sujatha, H. N. Ramaswamy and C. S. Yogananda), Science Publishers and CRC Press (2012) and private communication from Sitabhra Sinha.

to read them for the geometric ideas that inspired their forms is bound to be a much more difficult task than that of interpreting books of instructions for their building. The primary aim of this chapter is therefore just to take general stock of the material that exists and to look for pointers that might serve to guide future study; naturally, in places it may look like no more than a wish-list. It goes without saying that anything in the nature of a conclusion that may be turned up along the way must be considered tentative.

3.2 Measures and Numbers

It is customarily taught that three fundamental physical units or measures are necessary to apprehend quantitatively the world we live in, namely those in terms of which we quantify length, time duration and mass or weight. To these we have to add a zeroeth measure, that of the number of (discrete) objects in a collection or, in modern terminology, the cardinality of a finite set. A developed sense of numbers is a prior condition for the quantification of the other physical attributes; without it, there is no meaningful measurement possible. But even more importantly, counting in itself is the most fundamental and primitive measurement there is, not only of physical objects, but also of any discrete set of things we can imagine in our minds. Together with the study of shapes of objects in space, figures in the plane to begin with, numeracy is the essential starting point of all mathematical thinking. It is natural then to try to approach the question of what we can learn about the mathematical knowledge of the earliest phase of Indian civilisation, that of the Indus Valley people, from the angle of both numbers and measurement and of the geometry of shapes. It is a difficult task. Their writing has remained undeciphered and no oral tradition has preserved their spoken language; especially in assessing their level of numeracy, this is a severe handicap.

The task would have been somewhat easier if there was a discernible degree of continuity between the Indus culture and the Vedic culture which succeeded it, about whose mathematical attainments we have a very good idea. Indeed one motivation for looking at Indus Valley artefacts with an eye attuned to what came after is the possibility of finding such connections, even if tenuous. In the domain of geometry, the situation is not without hope for the future, as we will see. But the fact that a large part of this chapter is concerned with looking for such a linkage, any sort of linkage, is a measure of how little we actually know.

We have to presume that, like all ancient peoples, the Indus people also would have acquired an objective sense of the passage of time from the observation of natural periodicities like the solar day and year and the phases of the moon, even though there is no surviving evidence of any such astronomical observations (unlike in Mesopotamia). We also know nothing about the means adopted for the subdivision of the day into smaller time units. In the absence of narrative descriptions, the only way to get information of that kind will be to

discover instruments (clocks) specifically designed for the purpose. No artefact which might have been a clock has been identified, nothing like a sundial for example, though rudiments of shadow geometry as in the determination of the east-west line may very well have been known and understood.

In contrast to time, the spatial dimension, fortunately, is not ephemeral. Objects which are consciously designed and fabricated in durable materials are a valuable potential source of information about how lengths, areas and volumes were specified and measured and, more generally, how well geometry was understood and put to use. We may even hope to overcome the drawback of not having textual sources, at least to some extent and over time, by a painstaking study of the remains of the architecture – so systematic and complex at the same time – and its constituent elements. Such studies are still in their infancy.

It is evident that absolute numerical values of individual lengths and weights are not of much mathematical interest as the units used are arbitrary and, in the context of early civilisations generally, often given in terms of the human anatomy and hence variable. What matters are the relations that may hold among various dimensions, like linear or higher-dimensional proportionality or the Pythagorean property (as in Vedic structures), that override the variability in the actual physical lengths. (As a rule, modern archaeological writing supplies information on the size, weight, etc. of objects in modern units and that is useful in relating them to scales we are familiar with and in comparing artefacts from different cultures). In this respect the Indus people are exceptional in the regularity of the proportions that they observed in many aspects of their activities, as was already recorded by Marshall and his team of collaborators during the first extensive excavations.

Of these regularities, the most conspicuous example concerns the dimensions of the baked bricks. The use of high quality baked bricks began at or around the beginning of the mature phase (ca. 2600 BCE) and then was rapidly adopted in many city sites, most prominently by Harappa and Mohenjo Daro. The bricks are of an exceptionally high technical standard – attesting to a mastery of the technology of fire – and a majority of them were made using moulds whose lengths, widths and depths were in the ratio 4:2:1, a sequence found in other instances of spatial organisation such as " in the rooms of houses, in the overall plan of houses and in the construction of large public buildings" (Kenoyer, cited above). Partly because of their abundance, these rectangular bricks are among the most intensely studied of the Indus artefacts and there is statistical evidence that the small deviations observed can be accounted for by the changes in dimension caused by the firing and hence that the sides of the standard moulds themselves were in excellent binary proportions. It is also to be noted that the standard binary ratios occur commonly well before the widespread use of the firing technology.

Several Indus sites also made use of wedge-shaped bricks (in much smaller numbers, naturally) for the building of structures with a circular plan, such as water wells. These are bricks whose horizontal sections are isosceles (symmetric) trapezia and so are of interest, in view of their later use in the Vedic

culture in the making of circular altars (Chapter 2) as well as the fact that the all-important "great altar" (*mahāvedi*) itself had the form of a symmetric trapezium. Kenoyer (cited above) gives the dimensions relating to one well; it had an internal diameter of 1.2 meters and 36 bricks, each of length (the non-parallel sides) 26 cm, covered the circumference. A 36-sided polygon is a pretty good approximation to a circle, but before asking what we can learn from it, we have first to admire the feeling for geometry expressed in such a utilitarian undertaking as accessing subsoil water.

These wells belong to the early part of the mature Harappan phase and represent, along with the trapezoidal bricks utilised in their construction, the first practical attempt in India to reduce the properties of the circle to an associated regular polygon; so let us take a moment to recall the elementary relationship between the two. The perimeter and area of a regular polygon of n sides inscribed in a circle of radius R are $2nR\sin(\pi/n)$ and $nR^2\sin(\pi/n)\cos(\pi/n) = (1/2)nR^2\sin(2\pi/n)$ respectively. As $n \to \infty$, the sine tends to its argument, and the two expressions tend to the circumference and area of the circle. How good an approximation a regular n-gon is to its circumscribing circle is therefore determined by the first correction to θ in the power series expansion $\sin\theta = \theta - \theta^3/3! + \cdots$, i.e., by the ratio of the second term to the first term which is $\theta^2/6$. For $n = 36$, the errors in the perimeter and area are $(\pi/36)^2/6$ and $(2\pi/36)^2/6$, just under one part in 800 and 200 respectively. Since several circular wells of varying diameter have survived, it is not unreasonable to surmise that, just from the number and dimensions of the bricks, the Indus architects could have concluded that the circumference-diameter ratio is a constant and arrived at a good approximate value for this constant, i.e., π. As for the ratio of the area to the diameter, that would have required slightly less simple empirical methods and also led to a less accurate value of π.

Alternatively, if the circle is approximated by the polygon in which it is inscribed, its perimeter and area are $2nR\tan(\pi/n)$ and $nR^2\tan(\pi/n)$ and the attendant errors are of the same order.

We can in fact squeeze a little more information from the data. Assuming that the dimensions of the bricks have not been significantly eroded over time and that the gaps between adjacent bricks in the same course are small and uniform (so that the 36-gonal approximation can be trusted), we can conclude that the short parallel side of the trapezium (facing the centre of the circle) is the side of the 36-gon which, in the linear approximation to the sine and for an inner diameter of 120 cm, is $120\pi/36 = 10$ cm, for $\pi = 3$. This is the short side of the brick. From the length of the brick, 26 cm, the outer diameter is 172 cm and a similar calculation gives the long parallel side of the brick as 14.2 cm to an accuracy of better than a percent. Now 14 cm $\times$ 26cm is a common standard for rectangular bricks (the deviation from the 1:2 ratio seems not to worry the archaeologists unduly) and we can envisage the possibilty that the trapezium shaped moulds were made by starting with a rectangle and symmetrically trimming one of the short sides. What geometric method was used in this? More generally, how was the architectural geometry

involved in going from a circle of a given diameter to a regular polygon of a given number of sides and to the exact dimensions of the trapezium of each brick worked out? There must have been a systematic procedure; since the wells are not all of the same diameter, we can rule out the possibility that the dimensions were arrived at once (empirically?) and then frozen and passed on to the brick makers and builders. It is intriguing and fascinating that the most direct method of accomplishing the task, as we have seen above, is to rely on Aryabhatan trigonometry.

Symmetrically shortening one of the short sides of a rectangle is the preferred way to make isosceles trapezia (as for the great arena) in the *Śulbasūtra*. The numerical relationships among the three independent dimensions of the trapezium appear to be quite different from the Indus ratios, however. From the instructions, we know that both in the construction and the measurements of the great arena Pythagorean (diagonal) triples were made use of; the two parallel sides were drawn by making perpendiculars to the axis of symmetry (which becomes, eventually, the altitude of the trapezium) at its two ends and the length of the altitude and half the base (the long parallel side) were chosen as the smaller entries of an (integral) Pythagorean triple, no doubt to enable the use of one of the peg-and-cord methods for making right triangles. Apastamba in particular has a set of triples used for the purpose. For the well brick we have been looking at, the ratio of altitude to half-base is 26:7.1, which is about 3.7; none of the triples from Apastamba comes close to this in the corresponding ratio. From among 'small' triples, the nearest to these proportions is (25;7,24) in the notation of Chapter 2.2 which is not in Apastamba's list. It is difficult to ascribe an excess of 2 cm over the desired 24 to statistical variation or to natural causes (generally resulting in a contraction) even after we correct for the minor approximations made in the discussion above.

It would seem justified then to conclude that while binary ratios came to be an integral part of the Indus people's handling of linear dimensions, the Pythagorean paradigm, whether expressed numerically through triples, or geometrically, did not play a role. How then was perpendicularity implemented in the design of brick moulds, street plans, etc.? Very likely by resorting to another, non-Pythagorean, method of producing right angles that relies on the orthogonality of the common chord of intersecting circles and the line joining their centres. We will return to this question in the next section.

Even stronger evidence for the prevalence of binary sequences comes from weights. Once again, it is the beginning of the mature Harappan phase that marks a widespread standardisation of weights. Unlike in the case of lengths where only two fragments of graduated measuring rods have been identified, a large number of weights (and even a balance arm and pans) have been unearthed in excavations over the years. They are fabricated in rocks like chert and quartz, cubical in shape, and are polished and finished beautifully, almost jewel-like in appearance. The binary sequence dominates the lower end of the mass scale: taking the smallest weight as the unit, the sequence doubles it at each step until we get to 64. But, very interestingly, there is a second sequence which begins

to appear at 16 and this sequence is decimal: 160 and 1600; 320 and 3200; and 640 and 6400.

In drawing conclusions about the counting system underpinning the progression of the weights, the usual cautions apply, as in the case of linear measures. First, the inaccuracy of the ordered progression, whether binary or decimal, is about 5%, a little better when weights from the same site are compared. This may be due to any number of reasons: natural erosion of the stone, slow chemical processes over four millennia involving release or absorption of gases, lack of precision in the workmanship, etc. It would also be useful to know the ratios of the volumes of these stone weights or, equivalently, their densities; if the density is uniform across different weights, that could open the door to the possibility that the calibration might have been done by doubling volumes rather than by trial-and-error weighing, leading hence to a good idea (within at most 5%) of the value of the cube root of 2. In the absence of detailed and readily available data, such thoughts must remain idle speculation.

Overall, we may then take it, at least as a working hypothesis, that the Indus people counted using a base which is some power of 2, with a special role for 10. What power of 2? As it happens, there is a general consensus about the signs of the Indus script that represent numerals. These are groups of vertical strokes in one or two rows, up to four in a row (up to the number 8, that is) one below the other. In thus representing a numeral by the corresponding number of tokens (lines in the present case), the Indus enumeration is similar to other early civilisations. There are no representations with more than eight strokes (there is only one with eight) and that is consistent with the use of 8 as the base for counting. The interesting question is how numbers bigger than 8 were represented. If anything like a positional notation was used, the way to go about looking for an answer to the question would be to systematically search for symbols adjacent to the numerals up to 7 and hence might stand for a power of 8 (e.g., 13 will be represented by the symbols for 8 and 5 in juxtaposition). That also remains work for the future. And, finally, was there a special symbol for 10 and what special numerical role was 10 and its powers intended to play if 10 was not the base? It is useful to recall here that the Old Babylonian civilisation, while working with the base 60, had an independent symbol for 10 and that simplifies the reading of a number enormously – a group of 59 symbolic tokens is no more easy to comprehend quantitatively than a pile of 59 real objects. There are tantalising possibilities here. The numerical tablets of Old Babylonia, beginning ca. 1800 BCE, come after the end of the Mature Harappan phase. We have excellent evidence of close trade connections between Mesopotamia and the Indus Valley during the mature phase, perhaps the best established of long-distance contacts in the ancient world. In the currently very active field of the origins of Babylonian numbers, these contacts should perhaps be one of the factors considered.

Whatever the answers to these questions might turn out one day to be, there is one place where we can seek corroboration of the idea that the base was octal and that 10 had a special place in counting, if we take seriously the

suggestion of many scholars that there was a proto-Dravidian language that descended from the putative Indus language and which evolved into Old Tamil and, in due course, all other Dravidian languages. As was first pointed out by the Dravidian linguist Kamil Zvelebil[3], the verb root for 'count' in Dravidian languages is the same as the stem of the numerical noun 'eight' (*eṇ*). The Dravidian names of the numerals up to 7 have no such significance; in principle, number names for numerals up to the base, for any base, is a matter of free choice just as, in a written notation, number symbols up to the base are arbitrary, independently of the way composite numbers are subsequently formed. Even more suggestive is the fact that the name for 10 has no relation to any other numbers (it is certainly not the name for $8 + 2$) and the name for 9 translates literally as 'one-ten' whose only reasonable interpretation is as 'ten less one' in all four major Dravidian languages. Indeed, this rule scales decimally: 90 is 'hundred less ten' and 900 is 'thousand less hundred' in Tamil and Malayalam to this day (Telugu and Kannada, which were the earliest Dravidian languages to be Sanskritised, call them by the equivalent of nine times ten and nine times hundred). Clearly both 8 and 10 had special significance in the counting system in Old Tamil whose vestiges are still present in modern Dravidian languages.

There are other similarities between Indus archaeological data and near-modern Dravidian practices regarding numerical quantities. Of these, the most striking have to do with weights. In a paper which is difficult to read (and difficult of access), K. Venkatachalam[4] has compared the traditional weighing systems of south India (which survived the British rule, in fact until the introduction of the metric system post-independence) with the data from Harappa and Mohenjo Daro. Venkatachalam makes the point that the Tamil weight system which has been in use "from time immemorial to the recent past" is regulated by a combination of binary and decimal progressions. He also notes that many of the Tamil and Harappan weights are identical (within reasonable errors one might add) and that the unit weights in both systems are the same, about 0.87 grams, being 8 times the weight of a seed called *kunri-maṇi* in Tamil and *gunja* in north Indian languages (natural variations in the weight of the seed taken into account); for the Indus weights, Kenoyer (cited above) also mentions this possibility. It is an intriguing fact that, until decimal coinage came into force in the 1950s, Indian currency was binary: the rupee was divided into 2, 4, 8, 16 and 64 parts.

For many prehistorians and archaeologists, all these parallels are evidence supporting a past link between the Indus Valley culture and that of the deep south of India. The more exciting position for us to take (with caution as always) is to accept such a connection as tentatively established on the basis of non-numerical evidence alone. (After a critical analysis of the many abortive

[3]Cited by Walter A. Fairservis, "The Script of the Indus Valley Civilization", Scientific American, vol. **248** (1983), p. 58.

[4]K. Venkatachalam, "Weights and Measures of the Indus Valley Civilization", PILC Journal of Dravidic Studies, vol. **12** (2002), p. 119.

attempts at decipherment, Zvelebil says in a recent article that, of all the affinities of the Indus script with living languages suggested solely on the basis of linguistic comparisons, "the most probable candidate is some form of Dravidian"[5]). We can then turn the otherwise unexplained coincidences in numerals, weights, etc. around and use the Dravidian evidence to strengthen the case that the Indus people did count binarily, probably in base 8.

The obvious next task is one for the scholars who have recently made quite remarkable advances in the analysis of the syntactic structure of the Indus script. The Indus people presumably did not stop counting at 7. Can the symbol (or symbols) for 8 be identified definitively? What was the symbol for 10 and was 9 written in a way reflecting its spoken name (ten less one) in Old Tamil? More generally, how were composite numbers (those greater than 8) represented? Were there independent symbols for the powers of 8 or 10 and, if yes, do they occur in association with the well identified numerals up to 7? Given that the positional number notation of old Babylonia comes seven or eight centuries after the onset of the mature Harappan period, in fact coincides with its end, the importance of trying to find answers to such questions for the history of numeration cannot be exaggerated. The undertaking will not be easy, but it can be approached by statistical-analytic methods, without getting enmeshed in the almost hopeless complexities of any attempt at a decipherment of the whole language.

3.3 Geometry

Along with the macroscopic geometry of street alignment and architecture, it turns out that a most fruitful resource for the study of geometry is the vast inventory of practical and ornamental objects unearthed already in the 1920s and augmented over the years. These include objects like circular buttons, button seals and beads, as well as geometric designs, either incised or stamped on seals and pottery, or painted on jars, etc. While the large scale geometry is visually dominated by lines intersecting at right angles, geometry in the small as depicted on these objects is pervaded by circles, sometimes isolated, more often intersecting in regular patterns and generating thereby very striking decorative designs. It is best to begin by looking at artefacts in which the circular theme is evident or can be discerned more or less directly before going on to an identification of the geometric ideas they are based on. That will then lead on to the large scale constructions involving orthogonal lines in which the role of the circle is not immediately obvious. Along the way, I will raise the question of possible continuities between what we can learn from the Indus Valley objects and *Śulbasūtra* geometry. One striking link between the two will be identified;

[5]The work of Iravatham Mahadevan over the last two decades has significantly increased the probability. See Iravatham Mahadevan, "Dravidian Proof of the Indus Script via the Rig Veda: A Case Study", Bulletin of the Indus Research Centre, no. 4 (2014) and his other papers cited there.

more detailed study of Indus geometry, an activity still in infancy, will perhaps help find other connections

Of the various types of small objects based on the circle, the least interesting mathematically are the beads. The manufacture of beads out of semi-precious stones was a major industry starting from well before the mature Harappan phase. Most of them are cylindrical in shape; some have a relatively long axis, but the largest number are those called 'circular', with a length of the order of a millimeter or less and a diameter varying between roughly 15 mm and 1 mm. The perfect uniformity of the circular shape and the precision of the central hole make it likely that they were turned on some sort of lathe and then cut along the axis. Geometrically, they are of interest only to the extent that they testify to the fabricators' familiarity with the axial and central symmetry of the cylinder and the circle.

A little more interesting are patterns produced on pottery by incising or stamping circles on the wet clay. The circles are aligned along parallel rows, intersecting pairwise along each row (with no intersections between the rows), all of the same diameter but with the separation of the centres unrelated to the diameter. In the next step (not necessarily in chronological order; often, several types of decorated pottery were found in the same archaelogical stratum) the circles intersect in the plane – along both the 'X' and the 'Y' axes – giving rise to a pleasing pattern. In all of these designs, the circular impressions are created by stamping with a carefully made implement of some sort – the uniformity of the circular outline is as good as in the beads – though the intersections seem to have been guided by the eye rather than precisely positioned.

The next stage (again, not necessarily chronologically) of geometric development is seen in decorations painted on pottery such as the middle panel of Figure 3.1. Being painted by hand (on a curved surface), the design is executed somewhat roughly but there is little room for doubt that it is the realisation of a precisely conceptualised lattice. It is reasonable to think that the painted pattern was copied from an accurately drawn planar template; the template itself could not have been created without a good understanding of the properties of equally spaced intersecting circles of the same diameter. Because of its simplicity and symmetry, it is one which lends itself best to a geometric enquiry such as the one of Sinha cited earlier. Firstly, it is easy to identify a unit cell that will generate the whole lattice. The unit cell is not unique. The choice of Sinha *et al.* is shown in Figure 3.2. It is visually natural and there are objects like the ivory ornaments shown in Figure 3.3 which illustrate it in isolation. Keeping this in mind, the obvious way to construct the whole design is to draw two overlapping close-packed sets of equal circles, one displaced with respect to the other by half a lattice spacing along both axes (the solid and dashed circles in Figure 3.2). (Alternatively, we can think of the same pattern as resuting from circles whose centres are spaced $\sqrt{2}$ times the radius apart, along axes inclined to the X and Y axes by $\pi/4$). Sinha *et al.* then go on to show that other periodic patterns seen on the artefacts like the top and bottom bands on the jar of Figure 3.1 as well as a particularly striking 'tiling' by means

Figure 3.1: Painted jar from Mohenjo Daro

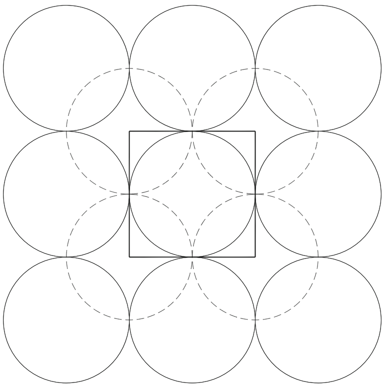

Figure 3.2: A choice of the unit cell for the lattice depicted on the jar.

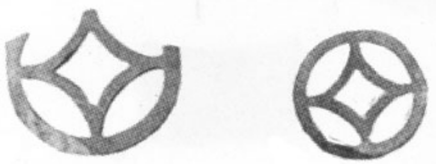

Figure 3.3: Ivory ornaments

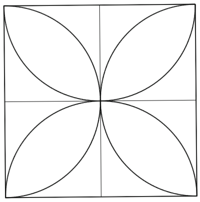

Figure 3.4: An alternative unit cell

of convex and concave circular arcs (see the paper cited above) can be derived from this basic pattern by deletion of suitable lines, shading, 'decoration', etc.

How were these lattices actually drawn? Let us first observe that an equally satisfactory choice of the unit cell is obtained by translating that of Figure 3.2 by half a lattice spacing (the common radius of the circles) either along the X-axis or along the Y-axis as shown in Figure 3.4. This is a design that we have met before, in Baudhayana's first construction of the square in the *Śulbasūtra* (Figure 2.7). It requires little imagination now to suggest that the whole lattice was generated by replicating that construction along both directions as many times as desired.

Here is one way to go about it. Begin by drawing the east-west line (or any straight line if orientation is not a consideration; in that case it just defines an axis), mark a point A_{00} on it and draw the north-south line through A_{00} by the *Śulbasūtra* method of intersecting arcs. If $2a$ is the desired side of the square unit cell, draw a circle of radius a centred at A_{00} and let it cut the east-west line (to the east) and the north-south line (to the south) at A_{10} and A_{01} respectively. Draw a series of circles, with centre A_{10} cutting the east-west line at A_{20}, with centre A_{20} cutting it at A_{30} and so on, all of radius a. Similarly draw circles with centre A_{01} cutting the north-south line at A_{02} and so on. The circles centred at A_{10} and A_{01} will intersect at A_{00} and also at A_{11}. Now draw the circle with centre A_{11} and let it intersect the circles centred at A_{20} and A_{02} at A_{21} and A_{12} respectively. Finally, A_{22} is the point of intersection of the circles centred at A_{21} and A_{12}. Thus the unit cell is the square $A_{00}A_{20}A_{22}A_{02}$; it is also the Baudhayana square, with A_{10}, A_{21}, A_{12} and A_{01} as the midpoints of the sides. Described in words the construction sounds complicated (my notation, slightly overelaborate perhaps, follows modern practice), but it really is a straightforward exercise with a pair of compasses in hand – see Figure 3.5. The only geometric element involved in the drawing is the circle and the only implement needed is a pair of compasses (perhaps a cord stretched between pins, the equivalent of the *rajju* of the *Śulbasūtra*); every labelled point in the figure is the intersection of a circle with one of the axes or of two circles. And, as noted in connection with Baudhayana's square, there is only one geometric principle used: the line joining the points of intersection of two equal circles cuts the line joining their centres orthogonally and equally. In particular, the theorem of the diagonal and Pythagorean triples have no role to play.

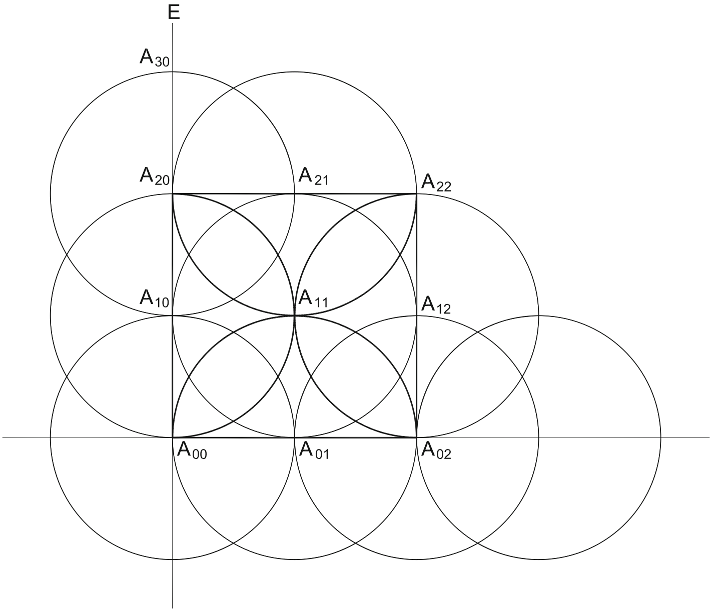

Figure 3.5: The '*Śulbasūtra*' lattice

On a macro-scale, a simpler version of this construction can be used to make the square and rectangular grids the street plans are based on. After marking the *EW* and *NS* axes, for the square grid it is sufficient to draw a series of circles of equal radii, centred on their successive points of intersection with the axes (so that they are equally spaced) and draw the common chords, extending them as far as necessary. Making the radii different for the circles along the two axes will result in a rectangular grid. We may speculate that it is in the making of the street plans that the orthogonality properties of intersecting circles were first put to use and that the decorative designs were a more evolved outcome of that geometry.

There are two different issues that have occupied us in this section: the geometric knowledge, by itself, of the Indus Valley people and the possibility that the geometry of the *Śulbasūtra* have elements which may have originated in the mature urban phase of the Indus culture. About the former, it is correct to say that the visible geometry can all be realised without resorting to the theorem of the diagonal and its arithmetical manifestation. In the absence of any evidence to the contrary, the prudent assumption will therefore be that the Indus artisans and town planners were unaware of the circle of ideas surrounding the theorem. In regard to the second issue, the fact that the steps in the construction of the square in the *Śulbasūtra* match so perfectly the Indus decorative patterns is a persuasive argument in favour of a causal link, but it is by no means decisive. Much work needs to be done to turn this tantalising glimpse

of a connection between the two into a reasonably well-established fact. I only note that in Baudhayana's *Śulbasūtra* – which is not only the earliest but also the one in which the organisation and order of presentation of the geometric material are the most logical – the construction of the square is the *first* (*sūtra* 4 of Chapter 1) technical topic broached, *before* the diagonal theorem is stated and illustrated with numerical examples.

3.4 Influences?

The absence so far of any evidence of a familiarity with Pythagorean methods in the Indus Valley civilisation is a good starting point for a quick first look at possible connections of Indian geometric traditions with other cultures, in particular the Old Babylonian. First, a recapitulation of the time lines. The sophisticated urban phase of the Indus civilisation and, along with it, the production of fine artefacts and the trade ties with west Asia, all came to a virtual end during 1900 - 1800 BCE. The Old Babylonian period covered roughly the three centuries beginning around 1800 BCE. It is to this time that the 'Pythagorean' tablets with their tables, diagrams and problems, all testifying to a knowledge of diagonal triples as well as the geometric theorem for squares belong. Given also the lack of evidence for Pythagorean thinking in the Indus remains (section 2 above), we have to conclude, in the present state of our knowledge, that the Mesopotamian developments were of indigenous origin.

The question now is whether this knowledge was somehow passed on to the Vedic Indians over the vast distances separating the loci of the two cultures. The situation here is ambiguous. The Vedic people began to occupy northwest India towards 1700 BCE before moving on to the Kuru country north of Delhi, settling there in the centuries preceding the formal composition of the earliest *Śulbasūtra*. Taking into account the scope and ambition of *Śulbasūtra* geometry – there is no comparable body of geometric knowledge contemporaneous with it – it is not unreasonable to suppose that it represents the finished form of ideas that had been in incubation over a considerable period, from the time the rules governing Vedic rituals were beginning to be set down. Direct archaeological evidence for a link between northwest and northern India and west Asia during this entire period, say 1500-1000 BCE, is very sparse if at all present, but there is an intriguing linguistic-mathematical connection. This comes in the form of a clay tablet from the Mitanni kingdom (on the border of modern Iraq and Syria, 14th century BCE) which has the names of the odd numerals 1 to 9, in a slightly modified but clearly recognisable Vedic Sanskrit but written in the Cuneiform script. Whether this is a pointer to a more extensive exchange of knowledge must remain a question for future research.

Equally frustrating is the lack of detailed information on what happened within the early Vedic heartland in the time span between the degradation and eventual disappearance of the larger Indus cities and the emergence of

the earliest Vedic texts. We do know from recent (and continuing) work[6] that the decline of the cities was accompanied and followed by the appearance of a large number of small towns and villages, unmistakably belonging to the late phase of the Indus culture but lacking any of the hallmarks of the urban phase at its apogee, in particular with none of the sophisticated artefacts or the high workmanship associated with the latter. This late 'small-town' Indus culture was spread over a largish area of north India, including the Vedic homeland of the time. Such a dispersal, especially if the de-urbanisation was provoked by environmental causes, is natural (and not in contradiction with the long-distance migration of a part of the population as postulated by the advocates of a proto-Dravidian - Indus connection). But, aside from a plausible Indus-Dravidian-Vedic linguistic link, we know next to nothing about possible interactions between the two peoples; and the little that we can guess is founded on imaginative interpretations of obscure allusions in the *Ṛgveda* or semi-mythical stories from the epics.

But knowledge lives on in the mind. It is not far-fetched to suppose that some of their architectural and geometric expertise lived on in the minds of the scattered Indus people, to be passed on to the Vedic altar makers around them – and it cannot be ruled out that they may themselves have been the altar makers – to become, in turn, the agent of a renewal. That is a less implausible scenario in any case than that the other agent of the renewal, the theorem of the diagonal, came to northern India from far away Babylonia.

The suggestion that there are elements of Vedic culture that show the influence of what scholars these days call the Indus ideology is not new. Generally, such suggested connections are at a fairly abstract level – for example, the supposed indigenous inspiration of the speculative philosophy of the late chapters I and X of the *Ṛgveda* – and impossible to subject to empirical examination. An exception is the link provided by the use of fired bricks in Vedic altars. As early as 1939, the great and celebrated Indologist Louis Renou touched on the mystery of the missing baked bricks in all Vedic constructions but the fire altars[7]. Subsequently, others have made a case, natural in the context, for not only the bricks used in the ritual altar but for the rite itself to be of non-Vedic provenance. It is undeniable that bricks do form a link between the urban Indus civilisation and the Veda (primarily *Yajurveda*; *Ṛgveda* has no mention of bricks) even if the other elements of the argument, involving philology and etymology, reconstructions of the migrations of both the Indus and the Vedic people – an end for one and a beginning for the other – etc., are not very compelling. There is, nevertheless, something to explain. Assuming that the

[6]Kavita Gangal, M. N. Vahia and R. Adhikari, "Spatio-temporal Analysis of the Indus urbanisation", Current Science, vol. **98** (2010), p. 846.

[7]In an article titled "*La maison védique*" Renou, after establishing that all secular building was done in perishable materials, concludes (Journal asiatique (1939), p. 481): ". . . Brick however is well-known in the *Yajurveda* but its use is limited to the piling up of the fire altar and, accessorily, to its annexes, . . . particularly singular when we consider the architectural technologies known to certain prehistoric civilisations in the northwest of India".

technology of fire was passed on by the residual Indus population, why were bricks not more widely employed and why only for making sacred structures? A plausible answer – the only one I know of – has two parts. The rural and pastoral Vedic people had no prior knowledge of the use of fire and well-designed kilns to bake clay into a hardy material and could not have embarked on the large-scale manufacture of bricks for general use on their own; they would have been dependent on the few among the indigenous people who still had the necessary skill. And it would have had the cachet of a new technology embodying mysterious powers and been surrounded by an aura of the magical and the supernatural. It would not have been the first time that mastery of a transformative force was followed by the bestowal of divinity on it – think of fire itself – nor would it be the last. The literature beginning already with the *Ṛgveda*, and the epics especially, are full of references to the supernatural skills of those beyond the Vedic pale in those arts and crafts, architecture in particular, that we associate with the Indus culture.

I conclude this chapter with a brief assessment of possible interactions between the geometrical cultures of Vedic India and ancient Greece, their similarities and disparities. For the latter, in order not to be overly dependent on speculation, it is prudent to rely on Euclid as representing the geometric ethos of the times leading up to him, due allowance being made for the logical rigour he brought to the organisation and presentation of the material. A similar role must be assigned to the *Śulbasūtra* as far as Vedic India is concerned; the scattered intimations of the geometry of ritual architecture in the Brahmanas and the still earlier *Taittirīyasaṃhitā* are too tenuous for anything like a coherent picture to emerge.

It used to be the case, beginning with the very first writings on the *Śulbasūtra* in the 19th century, that Vedic geometry was thought to be derived, as indeed all mathematics was thought to be, from the Greek. Many early European, especially British, scholars in fact placed the *Śulbasūtra* in the last centuries BCE, often for no better reason than to ensure the primacy of Euclid. The decipherment of the mathematical tablets of the Babylonians over the first half of the 20th century, together with their rather secure dating, made it impossible to hold on to the idea that the Greeks were the first mathematicians of the world. Much has been learned since then, not least of which, as far as India is concerned, being that we have now a credible and self-consistent chronology of the early Vedic texts. The *Śulbasūtra*, as part of the *Kalpasūtra* and carrying forward the unorganised mathematical ideas dating back to the *Taittirīyasaṃhitā* or earlier, fall within this period and comparisons, both linguistic and ritualistic, make it virtually impossible to make the earliest of them (Baudhayana) later than about 800 - 700 BCE. That rules out automatically any influence of Greek geometry on Indian. As for the possibilty of a reverse flow, there simply is no historical evidence for any sort of close contact between the two civilisations during the relevant period, say 1000 - 500 BCE. For those who would like to see an intimate relationship between the two traditions, there is then only one option left: that they evolved from a common source earlier

than about 1000 BCE and then went their independent ways, Greece a half millennium out of step with India.

The commonalities that Indian and Greek geometry share have been looked at with a degree of thoroughness by A. Seidenberg in a series of studies of which the last[8] is an accessible recapitulation. Seidenberg sees a parallel between ancient India and Greece in their association of geometry with ritual, though they are separated by about five centuries (with no proof of any contact during that time) and though the rituals themselves are quite distinct in inspiration and character. He also points to a common preoccupation with areas and area-preserving rectlinear transformation in the *Śulbasūtra* and Euclid (separated, again, by about half a millennium) Book II, quite a few of which are applications of the theorem of the diagonal. Against this has to placed the fact that the elements from which geometrical structures are built up are given distinctively different emphases in the two traditions. Books I and II of Euclid are all about lines and rectlinear figures, among which the triangle is preeminent. The circle is hardly present: it figures in four definitions (effectively one, the other three being definitions of the centre, the diameter and the semicircle) out of a total number of twentyfive, in one postulate out of five and in no proposition at all. (We have to wait till Book III for the first propositions on the circle). Parallel lines (and the parallel postulate) are all-important.

On the other side of the divide, the very first 'proposition' of the *Śulbasūtra* (Baudhayana) is about the use of circles and their properties for drawing squares with a given alignment and the alignment itself is determined by means of a circle. The notion of a pair of parallel lines is entirely absent, its place being taken by perpendicularity. Parallelism is a derivative and unremarked property possessed by two lines which happen to be perpendicular to a given line or, more generally, through their common orientation, e.g., lines in the east-west direction. Triangles are of secondary importance in the *Śulbasūtra*, occurring mainly as intermediate steps in certain constructions and, hence, in the shape of some of the prescribed bricks. And, of course, there are the overwhelmingly different foundations, both epistemic (Euclid's geometry is 'purer' than *Śulbasūtra*'s) and logical (there was no equivalent in India to Euclid's 'common notions', axioms of logic, at the time of the *Śulbasūtra*), upon which the two cultures built their geometries. All in all, despite the theorem of the diagonal and the area-preserving transformations – the universal truths of mathematics are no one's monopoly – the idea of geometry held in Greece and India was distinctively different one from the other, as was its practice.

If, despite these mathematical and epistemic dissimilarities, we wish to hold on to a linkage between Indian and Greek geometries, there are other difficulties, of a historical nature, to be overcome. It has been noted above that, given the constraints of chronology and history, the only viable scenario that will sustain such a linkage is for both traditions to have descended from an

[8]A. Seidenberg, "The Geometry of Vedic Rituals", in [Agni]. His earlier work can be traced from this reference.

earlier common source. This is the option favoured by Seidenberg. For him the obvious candidate for this postulated well-spring is Babylonia. But, as he himself says, “the dominant aspect of Old Babylonian mathematics is computational”. Among the thousands of mathematical tablets, there are few with a geometrical flavour to them and even those are concerned with what one may call numerical geometry, like the ones with the diagonal triples or the value of $\sqrt{2}$ inscribed along with the square and its diagonal. Standing somewhat apart from these however, though still connected to numerical area computations, is an unprovenanced tablet known as BM 15285 having some quite elaborate geometric figures.[9] Robson dates them to the Old Babylonian period, later than the urban Indus phase. Some of the figures involving sequences of circles are remarkable; one in particular is identical to the Indus Valley pattern generated by intersecting circles as in the middle panel on the jar of Figure 3.1. The tablet also has many figures involving squares with no circles in them, inscribed with a high degree of accuracy as far as the eye can tell from the reproductions. A few of them could well have served as illustrations of the cut-and-paste approach to the theorem of the diagonal for the square, similar to Figure 2.2, but not for the rectangle. Intriguing question: how were the squares constructed?

And, as long as we are willing to look at such tenuous parallels, here is one more. Among examples of the Greek association of geometric ideas with rituals noted by Seidenberg is the well-known legend in which the oracle of Delphi asked the citizens of Delos to solve the problem of doubling the volume of the cube as a means of warding off a plague (there are other versions of the story). The *Śulbasūtra* have nothing relating to the duplication of volumes, in fact nothing about the geometry of volumes. But, as we saw in section 2 of this chapter, the very mundane binary progression of Harappan weights required a solution of the same problem – that of determining the cube root of 2 – and it was achieved to a reasonably good approximation. That the duplication problem cannot be solved exactly through compass-and-ruler constructions, just as the *Śulbasūtra* problem of squaring the circle cannot be, is not relevant here. As far as I know, there is nothing about the cube root of 2 in Mesopotamian records.

It would seem on the whole that we are not yet close to a definitive appraisal of the interrelationships among Indus Valley, Mesopotamian, Vedic and Greek geometry. Chronologically and on the basis of one or two very faint indications, the least unlikely channel of communication, if there was one, is from Indus Valley to Mesopotamia to Greece. That there were commercial (and perhaps other) transactions between Mesopotamia and the coastal areas of the Harappan domain is an established fact. But as far as Vedic geometry itself is concerned, the problem of finding evidence of physical contacts between west Asia and north India during the relevant time, 1,500 - 1,000 BCE, remains unresolved. The search for mathematical similarities does not fare any better; without a detailed thematic working out of how either the ‘pure’ (largely

[9]See the reproductions in Eleanor Robson’s article “Mesopotamian Mathematics” in [MEMCII].

number-free) geometry of Euclid or the 'mixed' geometry of the *Śulbasūtra* could have evolved from that tradition, the hypothesis of a Mesopotamian influence on the *Śulbasūtra* in particular must be considered to remain, as of now, something to work on in the future.

For the present, it seems more realistic to think of Vedic geometry as the natural outgrowth of the once flourishing culture of the Indus people, close to hand and not fully extinguished, without a Babylonian mediation. Sometime along the way, the discovery of the Pythagorean paradigm, either locally or by infusion from another part of the world, transformed it into the mathematics that finally got written down in the *Śulbasūtra*[10].

[10] Seidenberg considers and rejects the possibility that both Greek and Vedic geometries are descended from the Harappan, preferring to attribute the source to the original Indo-Aryan 'invaders'. In the light of subsequent work by prehistorians, the latter possibilty can no longer be entertained.

4

Decimal Numbers

4.1 Background

Given the primordial importance of the understanding of numbers in the first awakening of a mathematical sensibility in early civilisations, it is something of a surprise that we have no text from India, comparable to the *Śulbasūtra* for geometry, devoted to the subject. Or perhaps it is not so surprising: it would seem that early people learned to use numbers in the natural course of events, without any conscious mathematical effort, just as we do today in childhood. The faculty of counting precisely and doing elementary arithmetic has become such a routine accomplishment that it is now almost part of basic literacy; some cognitive scientists in fact hold the view that it is innate. The fact remains that counting, i.e., the accurate determination of the cardinality of a finite set, and the manipulation of the numbers so determined are an acquired skill demanding logical thinking and a degree of abstraction.

The universal way of dealing with numbers today (by 'number' without qualification is meant, in this chapter, a positive integer), in counting and in arithmetic, is through the use of 10 as the base, i.e., by taking 10 as the standard or unit in terms of which any number larger than itself is expressed. The resulting decimal number system originated in India some time before the final compilation of the *Ṛgveda* (which will turn out to be an invaluable source of information on early numeracy in India), definitely before the 12th-11th century BCE. Historically, however, the general idea of using a base to manage numbers in a quantitatively precise fashion was first put to use – going by written evidence – during the Old Babylonian phase of the Mesopotamian civilisation, around 1800 BCE, with 60 as the base. As noted in connection with geometry (Chapter 3.4), the 19th century BCE marks two other significant developments, both in India: the beginning of the rapid dissipation of the Indus civilisation into an unsophisticated small-town culture and the ingress of the Vedic people into northwest India. So, once again, we are faced with the intriguing question

P. P. Divakaran, *The Mathematics of India*, Sources and Studies in the History of Mathematics and Physical Sciences, https://doi.org/10.1007/978-981-13-1774-3_4

of possible influences of the earlier cultures on the new arrival on the scene, not in the field of geometry this time, but in the practice of enumeration.

In the written notation of the Babylonians, numerals up to 9 were represented by what may be called literal symbols, the appropriate number of replicas of a token representing 1, e.g., a collection of three of them stood for the numeral 3. Such a literal approach to the problem of counting is bound to break down as soon as moderately large numbers are sought to be represented; to be presented with a set of n identical inscribed symbols in lieu of n concrete objects does not make the task of an enumerator any easier. The Babylonians solved this problem by inventing one symbol to stand for 10; thus 27 was represented by two copies of the '10' symbol and seven of the '1' symbol. This was a significant first step in the evolution of number symbols towards the abstract as it freed the choice of a symbol from any sensory association with the relevant cardinality, the value of the number. Numbers up to 59 were represented in this way but starting with 60, the now-familiar positional notation took over in which the value of a cluster of symbols representing a number up to 59 was decided by the place corresponding to the power of the base it occupied, $60^0 = 1$, 60, 60^2, etc., just as we do today with base 10.

The Indus Valley people also represented low numbers up to 7 (or possibly 8) literally, as collections of vertical strokes. But as was noted earlier, it is absolutely unknown what system of symbols they employed for larger numbers. While it is a plausible inference from the measures of length and weight and from words referring to numbers in Dravidian languages – assuming the validity of a link between Indus and early Dravidian – that they counted to base 8, there is also the special prominence given to decimal progression in the weights.

So, where did the fully-fledged decimal number nomenclature of the *Ṛgveda* come from? We are in the same situation here as in the case of geometry. If there was contact between Mesopotamia and the very early Vedic culture in India, it is not impossible that the idea of using a base could have come from there. Some historians in fact have suggested such an influence but, as we have seen before, the possibility of a contact between the two remains highly speculative. Moreover, as will become clear in the next chapter, there is no evidence whatsoever for a use of 60 as a number 'marker' in India at the relevant time. (The role of 60 in the subdivision of angles and time units came very late, relatively speaking, around the 5th century CE, and there is a perfectly good historical explanation for that). A possible scenario is that somewhere along the way it was recognised that using a special symbol for 10 but continuing with 60 as the base was an avoidable duplication. But the idea has no tangible support from either Mesopotamia or India; in particular, fractions continued to be written sexagesimally in Mesopotamia while they were expressed using decimal numbers in India from the beginning.

The situation regarding a possible Indus - Vedic continuity as regards numbers is hardly any better; there simply is not any sort of evidence for

a special place for 8 and its powers among the numbers in the *Ṛgveda*. A determined partisan of such a connection might point to the fact the Vedic name for 9 (*nava*) has also another meaning, 'new', that goes back to early Indo-European, as though its use to denote 9 came as an afterthought rather like the Dravidian 'ten less one'. For the present, it is futile to speculate; what is beyond doubt is that, on the evidence of the *Ṛgveda*, 10 was unambiguously the base the Vedic people worked with from the beginning. And it has retained that position throughout history and, in the last few centuries, throughout the world.

The evolution of decimal enumeration in India is not without its share of puzzles, however. When writing became widespread after the 3rd century BCE, it was natural that written symbolic representations of numbers gained currency, mainly in inscriptions in the Brahmi script on either side of the beginning of the common era (the book of Datta and Singh ([DS]) has a very detailed account). The organisation of the Brahmi numbers was not fully faithful to place-value, apparently somewhat out of touch with the very systematic decimal principles governing the names of numbers that were in continual use from Vedic times. (There is an explanation for this which we will come to in Chaper 5.3). It is only towards the end of the 6th century CE that we have the first inscription attesting to a strictly positional notation (a three digit date, all three digits being distinct). Within two centuries after that, the decimal notation spread to the new Mesopotamia, the Caliphate of Baghdad, where it was called the Hindu system of numbers and played a role in the rebirth of Mesopotamian astronomy. In time, the Hindu numbers were adopted by learned men (and traders) of the Mediterranean Islamic lands. It is in this way that decimal numbers and the principles they were based upon finally reached Europe under the name Arabic numbers, certainly by the beginning of the 13th century (through Leonardo of Pisa, more famously known as Fibonacci), perhaps a little earlier. Another four centuries were to pass before they began to be widely adopted: in the second half of the 17th century, Isaac Newton was still referring to them as "the new doctrine of numbers".

It is a striking fact that the great creative explosion of mathematical activity in Europe that began in early 17th century – with Descartes as the key instigator if we have to attach one name to it – was perfectly timed to absorb the practice of decimal enumeration and the ideas that go with it. As much as the rediscovery of Greek geometry and axiomatism contributed to it, so did the "new doctrine of numbers". Aside from the obvious advantages in doing arithmetic, turning it into a rule-based and algorithmic routine that almost anyone could learn, decimal numbers and arithmetical operations on them served as the model for algebraic expressions in one variable – polynomials, rational functions and power series – and their manipulation. (It is in the context of power series that Newton paid his tribute to the "new doctrine"). During this period of the exuberant growth of arithmetical and algebraic mathematics, almost all of it dependent on a structural understanding of decimal numbers, and until axiomatism caught up with it in the second half of the 19th century, Europe

paid little heed to logical niceties and what we now call rigour. The fact that it is only occasionally that we come across an acknowledgement in modern histories of mathematics of the power and influence of decimal numbers (despite Newton and, later, Laplace[1] who played a prominent role in decimalising mensuration in republican France) is perhaps to be attributed to the readiness and ease with which the idea was adopted and made its own by 17th and 18th century Europe.

Something roughly similar happened in India earlier. When Brahmagupta introduced the names of colours to denote unknowns in algebraically posed problems, he began by recollecting the rules of the standard arithmetical operations, extended to include quantities both positive and negative as well as the zero. A clear understanding of these rules was essential for basing his algebra on arithmetic, especially for his strongly algebraic treament of the linear and quadratic Diophantine problems. Five hundred years later Bhaskara II, who contributed significantly to the elucidation of the Diophantine equations of Aryabhata and Brahmagupta, wrote his treatise on algebra (*Bījagaṇita*) in which every algebraic step is modelled on the corresponding operation on numbers. But logical rigour was not the mathematical virtue that Bhaskara II sought above all else; that came afterwards, in the last of our three phases, as we shall see later. Though mathematicians always had good logical reasons for their assertions, the conscious and deliberate insistence on structured argumentation based on sound reasoning is a hallmark of the Nila school. In particular, there was a harking back to foundational matters concerning decimal numbers, with a special stress on their unboundedness and their property of succession, both of which were indispensable in Madhava's calculus. It was also recognised, explicitly for the first time, that a decimally expressed number is the value of a polynomial when the variable is set to the value 10 and the coefficients are restricted to be less than 10, which insight was then exploited to define a general polynomial in one variable.

Given all this, it seems best to start our discussion of the genesis and evolution of decimal numbers in India by first summarising the issues involved from a modern perspective. The evidence from the earliest textual sources for their mastery, in particular the specific traits imposed by orality on the expression of numbers through names, fits well within that framework. So do the more formal ideas in the work of the Nila school (see the flow chart in the introductory chapter) which, therefore, are also most conveniently dealt with in the present chapter.

[1]Bernard Bru in his Foreword to Patte's critical translation ([SiSi-P]), in French) of *Līlāvatī* and *Bījagaṇita* mentions the fact that, in his course of mathematics at the newly established École Normale in 1795, Laplace introduced decimal place-value numbers (together with the concept of zero) and arithmetic with the words: "The language philosophically the most perfect will be one in which one can express the greatest number of ideas by the smallest number of words possible." And that he concluded ('in an inaudible voice'): "These things appear simple and, in effect, they are; but it is in their very simplicity that their fruitfulness lies."

4.2 Numbers and Based Numbers

It was only in the second half of the 19th century that European mathematics took in hand a task that had been foreshadowed much earlier in the work of Descartes and a few of those who followed him, that of providing a formal logical basis to all mathematical sciences. From the vantage point of today, it appears that, although there were voices which called for formal clarity, the exhilaration of creating powerful new mathematics had deflected Europe from the straight and narrow path of logical soundness during the two centuries that followed; and a part of the responsibility for this must be attributed to the impact the widespread adoption of decimal numbers had on mathematics as a whole. When the endeavour of building a solid logical foundation finally got under way, its model was Classical Greek logic, as it already was for Descartes. The basic mathematical notion was – and still is – that of a set together with operations on (or, synonymously, transformations of) its elements or members which define functions or maps on the set, all subject to accepted principles of logic. Within this general framework, a set is considered to be characterised fully by the properties of the operations defined on them, and these are formulated in terms of some primitive notions whose (semantic) significance is not to be subjected to further analysis; an irreducible set of such 'self-evident' notions is taken to be part of the axioms. These are supplemented by the axioms of logic: rules, equally beyond questioning, for drawing valid conclusions. Since everything is formulated in abstract terms, the axiomatic approach is very powerful as a methodology, allowing great scope for generalisation: mathematical objects having different concrete origins and different practical significance could all be treated in a uniform way as long as they were the same as abstract sets and the defining set of operations on them obeyed the same formal rules. The different concrete mathematical realisations are then to be thought of as different representations of the same underlying abstract structure.

The purpose of giving this impressionistic and inadequate sketch of the axiomatic approach to mathematics is not only as an introduction to the axioms defining numbers. While it will be recognised that it is very close in spirit and method to Panini's structural description of Sanskrit grammar, the fact remains that the axiomatic philosophy never had an influential role in the Indian view of mathematics, a topic we shall return to in the last part of this book. Nevertheless, it will occasionally prove to be enlightening to look at particular aspects of Indian mathematics – an example being the nuances in the treatment of decimal numbers in the late text *Yuktibhāṣā* – from a more abstract standpoint than most Indian mathematicians might have approved of.

In the final form given to them by Peano (1889), the axioms defining the set **N** of all numbers fall into three types. The first type is very general and deals primarily with the logical characterisation of the equality of two numbers, an issue which Indian mathematicians would not have given any thought to. The other two types are more pertinent: i) how to characterise the intuitive notion that the defining property of numbers is that of having a unique successor

in a way acceptable to formal logic and ii) how to ensure that the set $\mathbf{N}$ so defined is precisely and uniquely the set of numbers. Both these types of axioms are important in the context of decimal numbers and are given below, again somewhat informally, grouping together the axioms of each type (rather than in the separated form in which they were given by Peano, and generally followed in textbooks since then).

1 (Succession). There is a function S on $\mathbf{N}$ which is one-to-one: $S(m)$ is an element of $\mathbf{N}$ for every element m and distinct elements of $\mathbf{N}$ have distinct successors, i.e., for any elements m and n of $\mathbf{N}$, $S(m) = S(n)$ implies that $m = n$. There is an element denoted 1 (and eventually identified with the number 1) which is not the successor of any element, i.e., there is no m such that $S(m) = 1$.

2 (Induction). If $\mathbf{N}'$ is a subset of $\mathbf{N}$ containing 1 and such that if m is an element of $\mathbf{N}'$, so is $S(m)$, then $\mathbf{N}' = \mathbf{N}$.

It is easy to see that the succession axiom captures all the properties that we generally associate with numbers, in particular that every number except 1 is the successor of a unique number, its predecessor, and that $\mathbf{N}$ is an infinite set. It is to be noted, for purposes of later comparison with the Indian approach to numerical infinity, that the latter conclusion does not depend on any arithmetical operations we may define on the set. In the axiomatic approach, in fact, the rules of arithmetic are to be derived from the operation of succession, by defining $m+1$ to be $S(m)$ and iterating this addition operation n times to get $m+n$. Similarly the product $m \times n$ is defined to be the n-fold iterate of the addition of m to itself. The binary operations of addition and multiplication are thus maps from the product set $\mathbf{N} \times \mathbf{N}$ into $\mathbf{N}$.

The definition of the other two elementary operations, subtraction and division, as inverses of addition and multiplication is slightly less direct. For $m > 1$, $m-1$ is clearly to be defined as the predecessor of m; similarly, as long as $n < m$, $m-n$ is defined as the nth predecessor of m. (The ordering symbols $>$ and $<$ are with respect to the order defined by succession and of course have their usual meanings). The most natural way to remove the restriction $n < m$ on subtraction is to introduce negative numbers axiomatically, i.e., to extend $\mathbf{N}$ to the set $\mathbf{Z}$ of all integers, which also brings in zero as a number, 0. The evolution of the idea of 'nothing' into the number 0 has had much attention paid to it by historians, much more than has the concept of negative numbers. Both these notions, especially that of negatives, have had a chequered history, more chequered in Europe – from the Greeks onwards until as late as the 19th century – than in India.

The division of m by n, similarly, can be defined as an operation on $\mathbf{N}$ as long as m is a multiple of n by a number. For general m and n, we can proceed as in subtraction and define division as a set operation by extending $\mathbf{N}$ to include positive fractions or, more generally, $\mathbf{Z}$ to the set $\mathbf{Q}$ of rationals. But there is also another way of thinking of division, using what is called the division algorithm. It is a consequence of the Peano axioms that, given m and $n < m$, there exist a unique number $q \leq m$ (the equality holding only for

$n = 1$), the quotient, and a unique number $r < n$, the remainder, such that $m = nq + r$. Obviously, if m is a multiple of n, $r = 0$; and by including 0 in $\mathbf{N}$ (see the next paragraph), we can dispense with the condition $m > n$ as well. Division can thus be thought of as an operation on the product set $\mathbf{N} \times \mathbf{N}$, i.e., as a map from $\mathbf{N} \times \mathbf{N}$ to itself. This 'division principle' is the basis of various algorithms for division as well as for extracting roots, and also makes clear that division is iterated subtraction, removing n as many times as possible from m, just as multiplication is iterated addition.

It is obvious that, as far as the structure given to $\mathbf{N}$ by the succession axiom is concerned, the distinguished (or first or smallest) element does not have to be identified with the number 1; any other number, positive or negative, will do. The choice of 1 (or, as is often done in modern times, 0) as that element is a practically convenient one, as it makes $\mathbf{N}$ identifiable with counting numbers and the results of arithmetical operations as defined above compatible with our empirical experience.

The induction axiom is quite different in the nature of its formulation and logical consequences. There are several apparently distinct but equivalent (within the framework of the rules of formal logic) ways of reformulating it, emphasising different aspects and adapted to drawing different consequences. One reformulation is to postulate that the result of iterating the succession map on 1 indefinitely is the set $\mathbf{N}$ omitting 1; that ensures that the axioms together define the set of numbers uniquely. Another is as follows. Suppose we have an infinite sequence of propositions P_m indexed by m whose validity (or 'truth' as often said) is to be established for all m. For that it is sufficient to show that i) the first proposition P_1 is true and ii) the validity of P_m implies the validity of $P_{m+1} = P_{S(m)}$. This restatement is obviously the logical basis of all inductive proofs. The equivalence of these reformulations follows from set theory; demonstrations can be found, at various levels of formality, in textbooks on the subject.

Suppose now that we are given two numbers N and $b \neq 1$ with $b \leq N$. Then, by the division principle, $N = N_1 b + n_0$ for a unique $N_1 < N$ and a unique $n_0 < b$. If $N_1 \geq b$, divide it again by b: $N_1 = N_2 b + n_1$, so that $N = (N_2 b + n_1)b + n_0 = N_2 b^2 + n_1 b + n_0$. Repeat the process until, for some k, N_k becomes less than b; call it n_k. We have then

$$N = n_k b^k + n_{k-1} b^{k-1} + \cdots + n_1 b + n_0$$

where all the coefficients n_i are less than b and some (n_k excepted) may vanish. The iterated division algorithm thus expresses every number as the value of a polynomial in one variable when the variable is fixed at the value b (the base) with coefficients less than the base. The ordered sequence $(n_k, n_{k-1}, \cdots, n_0)$ of coefficients determines and is determined by N uniquely, the latter because of the uniqueness part of the division principle. Thus, if b is fixed once and for all, we can economise on the notation by dropping any reference to it and representing N simply by the sequence of coefficients:

$$N = (n_k, n_{k-1}, \cdots, n_1, n_0)$$

and simplify further by omitting the commas and the brackets:

$$N = n_k n_{k-1} \cdots n_1 n_0$$

as long as it is understood that the expression on the right is a sequence and not the product. For the now-universal choice $b = 10$, this is exactly how we represent a number in writing, with the help of single symbols for the numbers up to 9 and including 0, each of them occupying the 'place of ones', the 'place of tens', etc. On these symbols when they occupy the place of ones (" in their natural position", to paraphrase *Yuktibhāṣā*, see below) the successor function is given by $S(0) = 1, \cdots, S(8) = 9$ and $S(9) =$ (1 in the place of tens and 0 in the place of ones) written, in conformity with our universal notation, as 10. On 10, S is defined by $S(10) = 1S(0) = 11$ and so on until $S(19) = 1S(9) = S(1)0 = 20$. The pattern continues in the obvious way through increasing powers of 10. The general rule is $S(n_k n_{k-1} \cdots n_1 n_0) = n_k n_{k-1} \cdots n_1 S(n_0)$ applied recursively, i.e., if $n_0 = 9$, the result is $n_k \cdots S(n_1)0$ and so on. These rules of decimal succession are no more and no less than the carry-over rules of primary-school arithmetic. It is to be noted that a written positional notation cannot work without a symbol for zero to indicate a position that is empty – the Old Babylonian method of leaving the appropriate places literally empty caused confusion, especially if the empty places occurred in the beginning as in 10, 100, etc.

These are completely elementary matters – we are after all dealing with an elementary subject – but in the Indian context in which numbers were defined only by their decimal place-value expressions, such questions had to be, and were, addressed with care.

For the sake of definiteness and because of its familiarity, we will stick to 10 as the default choice of base - it entails no loss of generality in theory, though there are important practical considerations to be met by a 'good' base. For the moment let us note only that the numbers 1 to 9 (more generally, numbers less than the base) have a special role, in that all numbers can be built out of them and 0, by distributing them among the places. This property of the 'low' numbers was made much of in India by scholars in disciplines other than mathematics, the most perceptive (and celebrated) being the philosopher of language Bhartrhari (Bhartṛhari, 5th - 6th centuries CE probably), a near contemporary of Aryabhata. There are several places in his work *Vākyapadīya* ("Of Sentences and Words") where he alludes to the special role of "numbers which occur first" (*ādyasaṃkhyā*) as a model for the building up of sentences from words, for example: "Just as by means of the grasping of numbers which occur first (or in the beginning) is understood other numbers and the distinctions [among them], so [it is] with the hearing of words and their differences". (Note that words are heard, not read). Any lingering doubt about what he meant is dissipated by the gloss (possibly by himself) which goes: "As the numbers one, etc. $\cdots$ serving different purposes are the means of understanding numbers like hundred, thousand, etc., and are considered a constituent part of hundred etc., so the apprehension of a sentence is based on the precise meaning of words such as Devadatta (the name Devadatta is the normative noun in the nominative case

in classical Sanskrit writing on language and philosophy) etc., the understanding of which is innate". Taking the analogy a step further, he speaks elsewhere in the same text of the atoms (*aṇu* and *paramāṇu*) of word (or speech or sound) gathering together, by the manifestation of their own capacity, as clouds are formed by material atoms. So, in the Indian view, numbers up to 9 are like atoms, gathering together to form numbers larger than 10 in their infinitude. Consistent with Bhartrhari's simile and as an acknowledgement of his insight, it seems right to distinguish the numbers up to 9 (numbers less than b for a general base) as atomic numbers in terms of which all other (compound) numbers can be expressed, a nomenclature that will be followed from now on; it is less ambiguous and more accurately evocative than other usual terms like digits or numerals.

Indian philosophy had been concerned with questions of the logical foundations of methods of validation of statements ('propositions') since early and had expressed doubts about axiomatic methods at least from the time of the later Upaniṣads, ca. 6th - 5th centuries BCE (I will take up this issue in Part IV). Mathematicians and astronomers rejected such an approach at all times, most explicitly in the writings of Nilakantha in the early 16th century. It is therefore with slight surprise that we find Nilakantha's disciple Jyeshthadeva, in the first chapter of his book *Yuktibhāṣā* (written in the 1520s), devoting time to a description of numbers and enumeration that emphasises their property of succession and relating it to elementary arithmetical operations. The unexpectedness of this passage alone makes it worth quoting in full. The opening paragraph of the book sets forth the basics of decimal place-value enumeration starting with a characterisation of atomic numbers as *prakṛti*, 'in their original form or natural state', and then names the successive powers of 10 and describes their places, including the idea of carry-over. After that Jyeshthadeva introduces the fundamental operations:

> Now [I] demonstrate the different computations with these [decimal numbers]. All computations are of two types: of increasing nature or of diminishing nature. Computations which result in increase are addition, multiplication, square and cube. Those which result in decrease are subtraction, division, square root and cube root. Addition is of use in multiplication, multiplication in squaring and squaring in cubing. In the same way, subtraction is of use in division, division in [taking] the square root and the square root in [taking] the cube root. Thus, the preceding [operation] is of use in the succeeding.

What follows immediately is much more interesting:

> Now [I] show the method of carrying out these [operations]. When units (the word is *rūpam*, often used for 1 as the unit of counting) are added successively to a number, they (the results) will be the higher and higher numbers starting with it (the starting number). If, from a large number, units are removed successively, the results will be smaller and smaller numbers starting with it. In this manner

> all numbers get their individual identity (*svarūpam*, 'self-form'). If [we] recall all numbers higher, one by one, than any chosen number, [we] get the result of adding one successively to that number. (A similar statement about subtraction follows). So, by recalling the identities of increasing and decreasing numbers, [we] get the results of addition and subtraction to each number. Thus add separately ("in a different place") as many ones as [we] intend to add to a chosen number; add that many to the chosen number in one operation. The result will be the same number as that of adding the ones separately. That will be evident if [we] reflect on it. (A similar statement about subtraction follows). Therefore if [we] know how to count forwards and backwards, we get the results of addition and subtraction.

The first surprise here is that the issues around what is meant by addition and subtraction were at all deemed important enough to be given so much care and attention, 2500 years after decimal numbers (in the *Ṛgveda*) and arithmetic with them (in the *Śulbasūtra*) became part of mathematics. No other text in this long interval, at least none of the standard expositions, makes such an effort. It is customary for Indian mathematicians to pay their respects to knowledge already acquired, partly as a recollection of prerequisites on which the new work is to be based and partly as tribute to the great masters of the past. The number system generally gets a passing mention at the beginning of most works. To take one example out of many, Bhaskara II begins *Līlāvatī* by listing the names of the powers of 10 and also, very briefly, sets out the operational rules for decimal addition and subtraction (before doing the same for the other operations). But it is just a routine remembering of names and rules; there is no philosophy there, nothing about foundations.

The second surprise is that the recapitulation is done in such abstract terms. The conclusion is inescapable that *Yuktibhāṣā* in this passage is not merely summarising known material but urging on its students a new way of looking at numbers, in fact looking at them, for the first time, with a sharp logical eye. And, despite the scepticism that unquestioned axioms and 'first principles' always aroused in India, where the logic leads to is the almost Peano-esque realisation that numbers are primarily defined by their property of succession. It is a typical feature of *Yuktibhāṣā*'s style of writing to introduce innovative ideas or techniques, as in the passage above, by resorting to the time-tested pedagogic device of repetition, doubly reinforced here by the exhortation to "reflect on it" – the mental faculty that is called upon is not the memorisation of something axiomatic and hence beyond logic, but the capacity to think. The issue of the "nature of numbers" (the heading of the first chapter) is eventually settled in a pragmatic way, without appeal to formal rigour as we understand it today, but that cannot hide the resemblance of the approach to the logical impulses of late 19th century Europe. Though there is nothing in this part of *Yuktibhāṣā* that overtly connects it to the problem of induction, the identification of the nth successor with the result of adding n serves the same purpose as the induction axiom, that of asserting that the set of natural numbers is

unique. Despite this absence, I think it certain that we owe the preoccupation with the foundations of the notion of numbers chiefly to the logical demands made by the first appearance of mathematical induction as a method of proof.

In the context of these formal considerations, it is helpful for an appreciation of the Indian perspective on such questions to reemphasise the philosophical distinction between the Peano view and the decimal view of numbers. While they are both recursive in nature, the operation of succession is elementary, a matter of definition and logic, in the sense that it is a map from $\mathbf{N}$ to $\mathbf{N}$ with one number, which is designated as the successor, as the output on which the operation is iterated; the axioms then ensure that $\mathbf{N}$ is a set that is unique and unbounded above. In the decimal view, the 'existence' and unboundedness of numbers is not an issue. The division algorithm supplies an arithmetically constructed map from $\mathbf{N} \times \mathbf{N}$ to itself – it presupposes not only the existence of numbers in some intuitive sense but also a familiarity with elementary arithmetic – and the operation is iterated on one of the two output numbers, namely the quotient. The infinitude of numbers came to be thought of primarily as arising not additively through succession but from promoting an atomic number to a higher place by multiplication by powers of 10; the milestones on the road to infinity are exponentially spaced rather than linearly.

The recursive nature of the construction of decimal place-value numbers came to have a powerful influence on many aspects of Indian thought, most directly on mathematics itself. For our purpose, it is adequate to define a recursive mathematical process very generally as one in which a result of a particular operation (or, sometimes, a sequence of operations) is subjected to the same operation again, *ad infinitum* or until it terminates. Techniques based on the idea, in one form or another, are to be found in all areas of Indian mathematics during all periods, and the decimal number system is their mathematical ancestor and model. Several examples – as well as some formal characterisations – will be met with in the following chapters (especially Chapters 10.4 and 15.3); particularly relevant is a recursive procedure closely analogous to the place-value algorithm, the Euclidean algorithm, that was known first to the Greeks chiefly as a geometrical operation but was always part of arithmetic in India. In its finite variant, it is the main ingredient in the solution of Diophantine problems.

4.3 The Place-value Principle and its Realisations

This being a printed book, we are reliant on written words and symbols as the medium of discourse, even when dealing with early cultures whose mathematics was not expressed or communicated in writing. At the level of ideas however, in the mind as it were, we have to assume that the first conceptualisation of numbers and their properties followed similar paths in early societies, whether they were then put down in written symbols as in Old Babylonia or pronounced as names in Sanskrit as in Vedic India. Thus, though the representation of a

number in terms of its decimal coefficients was presented in the previous section as the written identification of N with the sequence $(n_k, \cdots, n_0)$, it has to be understood as a concrete representation, familiar to us from long usage, of an abstract idea that transcends the particular concrete manner in which it is presented and that can accommodate other modes of expression as well. It is therefore useful to inject a degree of abstraction in any attempt to reconstruct the thought processes of early civilisations as they came to terms with the ideas and processes that led to place-value enumeration, each in its own distinct way. The abstract substructure must include elementary arithmetic upto and including division with remainder, a knowledge of which, as we have seen, is a necessary precondition for the genesis of place-value counting – counting and arithmetic must have evolved in tandem, growing out of empirical attempts at precise measurements of discrete quantity.

There seems to be a consensus among historians that the earliest manifestation of the impulse to count took the form of making a one-to-one correspondence (matching) between two sets of objects, allowing a determination of whether they had the same number of objects or, if not, which one had the greater number. There is evidence in the early Vedic literature of India that such matching methods may in fact have been used. It is not difficult to imagine that the limitations of determining comparative cardinality qualitatively in this fashion would have been recognised and better methods sought: fix a number, 10, compare all numbers with it, if necessary repeatedly; in other words, use 10 as the standard or unit (the base, in modern terminology) to measure other numbers with.

Translating this operational description into more abstract language, I will use the term 'place-value principle' to mean the rules that allow us to measure (and hence to define), through a recursive arithmetical procedure, all numbers larger than the base in terms of any chosen base. The procedure is the one described in the previous section, that of finding the remainder after successive division of the quotients by the base. At this stage, the formulation of the principle and its implementation through the recursive use of the division algorithm are to be thought of as an inherent characterisation of based numbers, independent of the way it is expressed by symbols or names (or conceivably by other means), and of course also of the choice of a base; they constitute different realisations or representations of the abstract, unique, arithmetically defined object which is the set of plac-value numbers. Stated in this way, the principle reduces the precise understanding of a number, no matter how large, to the cognition of a small set of atomic numbers. It has a converse or complementary aspect, that of constructing an arbitrary number (bigger than the base) from an ordered sequence of atomic numbers. Here the relevant rules (or algorithm) are those for the formation of polynomials: take increasing powers of the base, multiply each such power by the atomic number which occupies that place and finally add all such multiples together. This is the process by which we 'read' numbers.

It is immediately clear that the principle is of no help in apprehending numbers smaller than the base or in doing arithmetic with them as there are no rules to apply – 'counting' atomic numbers is a different exercise from that of counting big numbers. How then does one internalise the fourness of 4 or the nineness of 9? The only reasonable answer that anyone has come up with is that we humans (at least) have an innate faculty for grasping the precise numerical significance of the cardinality of small sets. In support of such a view we can point to a fact mentioned earlier, that of the use of a literal symbolic representation of small numbers by the corresponding number of tokens in several early cultures. The evolution into the association of an atomic number with some general arbitrary symbol not literally and graphically related to the cardinality is a significant step since it recognises that any symbol can be used to denote any atomic number as long as the association of symbol to number is one-to-one and fixed once and for all and for all people. The mental faculty involved in making possible this transition to cognition of abstract symbols for the atomic numbers is that of learning and memory: which symbol represents what cardinality or, equivalently, what is the order in which the sequence of symbols is ordained to occur? In contrast, the cognition of large numbers results from an algorithm, no matter that it is performed almost automatically, and hence based on reasoning (as well as learning, but this time of the rules governing the algorithm). Research into the cognition of numbers has seen a great deal of theoretical and experimental activity in recent times, but much of it seems to be unmindful (unlike Bhartrhari) of the distinction between the way we 'take in' small and large numbers.

A very similar distinction can be drawn between the operations of elementary arithmetic involving atomic and compound numbers. The performance of such operations on large numbers is done by the application of certain rules and these rules derive their validity from the representation of numbers as polynomials in the base b. The rules of long addition and long multiplication for instance are precisely the rules for defining the same operations on polynomials together with the carry-over rules, the latter reflecting the succession rules for based numbers (as described in the previous section). Thus, in base 10 and in painful detail, $15+16 = (1+1)\times 10+(5+6) = 20+11 = (20+10)+1 = 31$ and $15\times 16 = 10^2+(5+6)\times 10+5\times 6 = 100+110+30 = (100+100)+(10+30) = 240$. The point of this little exercise is to stress that the rules can be applied only after we know the results of adding and multiplying atomic numbers for which there are no rules; sums and products of atomic numbers are to be arrived at by experimentation, physically or in our minds with the help of the counting rules, and retained in our memory. There is no substitute for the memorisation of the multiplication table unless it is the consultation of a written table once it is established for the first time – how else but empirically?

We can think (and people in the past have thought) of many different methods of expressing or representing the abstract place-value system of numbers and, in all of them, the association of an atomic number to whatever represents it – visual symbol, sound pattern, gesture, or any other – is, necessarily,

arbitrary (as long as it becomes, eventually, unanimous). To take up the now universally practised written representation first, the customary symbols 1 to 9 (not to forget 0) have their origin in historical events and their association with given cardinalities is purely conventional. The fact that they are simple visual patterns, easy to write, memorise and read, played only a relatively minor role in the overwhelming dominance of the written representation in modern times. The widespread acceptance of the symbolic representation came with the spread of writing itself and especially after the invention of printing with movable types; the sequence of symbols, written or printed in a line, fits perfectly into the linear structure of written and printed text, and it leaves a permanent record behind. There are other, more intrinsic, advantages as well. The demands on memory are minimal: nine symbols in sequence plus one for zero. The length of a symbolically written number scales more or less as the logarithm of its magnitude and is an economical rough measure of it: we know instantly that a number 10 places long lies between 10^9 and 10^{10}. It is totally without ambiguity, the only freedom being whether increasing powers are to be written from right to left or the other way. These points are so familiar to us that the only useful purpose of bringing them up may be to serve as a touchstone for the representation that we are most concerned with, the oral or nominal representation invented and perfected in pre-Vedic India.

In an orally literate (and numerate, with base 10) culture such as Vedic India, numbers have to be given their identities by creating distinct names for them, as indeed for every communicable idea – objects, actions, attributes, and so on. The designation of a number by a name need not be unique (as is also true in principle, though not generally in practice, in the case of symbols) but the reverse association must be; the same name cannot correspond to two different numbers. The naming begins with the atomic numbers and their names are, exactly as for their symbolic counterparts, arbitrary. These have then to be committed directly to the individual and collective memory, permanently, to be retrieved at will. The names of the numbers from 1 to 9 in the *Ṛgveda* are in fact unique as nouns: *eka, dvi, tri, catur, pañca, ṣaṭ, sapta, aṣṭa, nava* (their declensions when employed as adjectives can depend on the gender, number and case of the qualified noun). But. as we shall see in some detail in the next chapter, this is far from true for compound numbers. Depending on the grammatical rules utilised to combine atomic number names to produce large numbers, i.e., to linguistically implement the necessary arithmetical operations, the same compound number often has different linguistic expressions. Each such variant has to be analysed to find the exact numerical significance of the name and, in all of the many occurrences of such names in the *Ṛgveda*, there are perhaps only two or three instance of a grammatically defined name whose analysis can be made to yield more than one number. (This is not as trivial a matter as may appear: Panini's systematisation of grammar came six or seven centuries later). For the atomic numbers, this kind of potential ambiguity is not an issue since their names are primitive and no arithmetic or grammar is

involved. Indeed they also seem not to have had any prior semantic significance at all (except *nava* = new, as was noted earlier).

Once the atomic number names are fixed, the logic of the place-value construction of numbers requires a means of designating the places as the next step. Inevitably, this is accomplished by giving names to the powers of 10. The name of any given power 10^i could have been constructed by reference to the exponent i – much as in the use of *karaṇi* in the *Śulbasūtra* for the square root, *dvikaraṇi* $= \sqrt{2}$, the producer of twice (the area), *trikaraṇi* $= \sqrt{3}$ – but was not, probably because the general idea of an exponent was absent at the time the names of the first few powers were settled. The preferred option was to invent arbitrary names for the powers of 10 from *daśa* = 10 onwards as higher and higher powers came into common use. Some of these names already had other meanings but logically, as names of numbers, they did not need to have any relation to their literary meanings.

Given the names of the atomic numbers and the powers of 10, there is no more freedom left in the choice of names in an ideal method of nomenclature. By ideal is meant, primarily, that it must be rule-based and that the rules, which are the rules of grammar, must be general or universal. Thus, in the next logical step, that of combining the names of n_i and 10^i to get a name for their product, the rule should not depend on i. With such a multiplicative rule, we have in hand a name for each term in the polynomial representation of any number. Finally, the terms have to be added together, also by means of a grammatical rule of universal applicability, to arrive at a name whose only possible interpretation is as the number $\sum_i n_i 10^i = N$.

It is to be noted that the system of number names formed in this manner has no need for a name for 0 except when it occurs as a number by itself, to denote 'nothing'. We have already seen that the method of writing numbers as a sequence of the coefficients in its polynomial representation (or, equivalently, of the remainders after successive division by powers of 10) cannot do without a symbol for 0, though one can dispense with it if the full polynomial expression is written; number names are no more than a verbal expression of the corresponding polynomials.

In summary then, the systematics of number nomenclature needs an essential linguistic dimension to be added to its logical structure. The generic equivalent of addition is conjunction ('and' or synonymous terms) and that of multiplication is adjectival or adverbal qualification arising from a repetition ('times') of the performance of addition. Such grammatical operations are often implemented in Sanskrit by the process of nominal composition (instead of describing the process in words), i.e., by composing adjective and noun together into a compound word (*samāsa*) and applying the rules of *sandhi* at the junctions. The rules of composition and of the analysis of a compound word into its more elementary constituent words can be quite intricate and it requires an understanding of the logic of numbers as well as of Sanskrit grammar to decipher how they are used to arrive at a precise nominal designation of numbers. This merging of normally unrelated disciplines may well have been a contributory

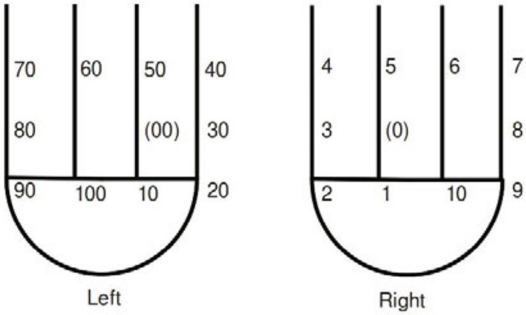

Figure 4.1: Gestural counting in the chanting of the *gāyatrī*

factor in the relative neglect of the Vedic corpus as the primary source for the early history of decimal enumeration.

4.4 Other Realisations

There are other realisations of decimal enumeration to be found in the Indian tradition and in Indian texts. Of these, the oldest, perhaps as old as the *Ṛgveda* itself, is closely related to a particular ritual and has survived to the present. Specifically, it is a method for keeping count in the recitation of a short *mantra*, the *gāyatrī*, required to be repeated 108 times (and, on special occasions, 1008 times). The *mantra* occurs in Book III of the *Ṛgveda* which scholars take as belonging to its earliest phase. The reciter keeps track of the count by what can be called a gestural place-value method, by associating some of the finger joints of the two hands to the atomic numbers (as shown in the schematic Figure 4.1), with the right hand as the place of ones and the left as the place of tens. The right thumb (not shown in the figure) moves over the joints of the right hand, starting at the position marked (0) and taking one step to the next numeral for each repetition. When it reaches 10, the left thumb is also shifted to 10 from (00); the right continues on to 1, 2, etc. until it gets to 10 again at which point the left thumb is moved to 20, and so on. On the completion of 100 counts, the reciter takes a blade of grass from the 10 he started out with and lays it to one side, the left thumb reverts to its starting position, and he carries on as before until all 10 blades of grass are shifted.

In the general context of the genesis of decimal counting, the main interest of this gestural method lies in its possible antiquity. If the ritual is indeed as old as the *mantra*, we have here a practice that is contemporaneous with verbal counting, thereby indicating a flexibility, dating from the earliest hisorical times, in the means adopted for the practical implementation of the principles of decimal enumeration. Though limited in its range by the finiteness at hand as it were, it is a more abstract realisation than that of giving names (which in any case could not be uttered while chanting the *mantra*) as it is independent of a supporting linguistic structure. Not having writing, or considering it a profane

skill, it is the closest the Vedic people got to a representation of decimals that is entirely free of the influence of the (spoken) language.

A much later (after the advent of written decimal numbers) and more versatile realisation, symbolic but still oral in inspiration, overcomes many of the limitations of the literal number-naming paradigm of Vedic literature. Called *kaṭapayādi* ('*ka, ṭa, pa, ya* and so on'), it is a syllabo-numeric representation that is almost completely abstract, with language having no role to play except that of supplying syllables. The central idea is to map each atomic number, including 0, to a set of Sanskrit syllables, taking advantage of the systematic organisation of the consonants (Chapter 0.4), in particular the division of the first twentyfive of them into five groups, (the columns of the consonant square) beginning respectively with *k, c, ṭ, t* and *p*, each consisting of five phonetically related consonants (the rows), e.g., *k, kh, g, gh, ṅ*. The ten consonants of the first two rows are associated to the numbers $1, 2, \cdots, 9, 0$ in order, independently of their vowel values, as are the consonants of the third and fourth rows. For example, *ka, kā, ki* etc. all get the value 1, *ca, cā, ci* etc. get the value 6 and so on. The consonants in the fifth and last row are assigned the numbers 1 to 5 and the consonants outside the grid, $y, \cdots, h$, are assigned the numbers 1 to 8. So, at this stage, we have 1 represented by any of *k, ṭ, p, y* (which is the reason for the name *kaṭapayādi*), with any possible vowel ending, and similarly for all numbers up to 8; 9 and 0 are represented by three consonants each. The system is now augmented by assigning to every vowel occurring by itself, i.e., not as the voiced ending of a consonant, the value 0. A final rule is that, in compound consonants, only the last consonant has a numerical value; thus *rā*, *tri* and *strī* all stand for 2.

The *kaṭapayādi* representation of a decimally expressed number is now obtained by replacing each of its atomic entries by any of the syllables, a few dozen of them on the average, to which it can be mapped. Every sequence of syllables, semantically sensible or not, in particular every word or phrase, thus gets a precise numerical significance. It is understood by convention that the syllable to the extreme left is in the place of ones, the syllable to its right is in the place of tens and so on; a word is therefore read as a number in the sense opposite to that in which its numerical expression is written (for instance *kaṭapayādi* $= 81111$, *saṃkhyā* $= 17$). There are two sources of ambiguity; the first, not really serious in practice, is to know where to supply breaks - is a given verbal expression to be read as one number or several shorter ones in sequence? The second is inherent in the fact that every sequence of syllables has a numerical value; there is nothing to stop a reader from picking up a phrase in plaintext and interpreting it, say, as the date on which the work was completed (and I believe it has happened; see Chapter 9.2 for a possible illustration). In actual usage, its greatest appeal to authors came from the one-to-many nature of the map $\{\text{atomic numbers}\} \to \{\text{syllables}\}$: it was easy for a skilled practitioner to convert any number into a meaningful Sanskrit word and there are instances of lists of numbers being turned into reasonably sensible phrases and even whole verses, which is of great help in memorisation.

The *kaṭapayādi* system is equally efficaceous whether it is used in spoken or written form. It shares with the symbolic written representation the advantage of being economical, the syllabic length of a phrase representing a number being the same as its place-value length. It is very likely that it originated as a response to the problem of dealing with long lists of often very large numbers arising in astronomy, e.g., planetary parameters or the table of sines, especially when the text is in metrically regulated verse. This need was felt already by Aryabhata who invented an even more compact syllabo-numeric notation - the largest number so expressed by him (in five syllables) is 1,582,237,500, the number of cycles ('revolutions') of the moon in an astronomical epoch or *yuga* – but at the cost of not respecting the decimal structure. Perhaps for this reason or because of its artificiality, the Aryabhatan code did not become very popular. We see the fully formed *kaṭapayādi* notation, in great profusion, for the first time in an astronomical manual of Haridatta dated 684 CE, less than two centuries after the writing of the *Āryabhaṭīya*. The system became a great favourite of several authors of the Nila school and it has been suggested that Haridatta, whose claim to fame rests on an updating of the astronomical parameters as given in *Āryabhaṭīya*, was in fact an inhabitant of Kerala.

From a theoretical point of view, one noteworthy aspect of *kaṭapayādi* is its Paninian abstractness: denoting a whole set of syllables by one number which determines how the rules of arithmetic apply to all the syllables representing that number, a sort of algebra-in-reverse. The fact that it lends itself so readily to the memorisation of long lists of large numbers through literary devices is an added bonus, doubtless of value to the teachers and students of the time. It was not an outgrowth of the purely oral tradition, more likely a method of transcribing numbers which already had a written decimal existence, especially those whose names are very long, into text. For purposes of enumeration it is as powerful as the standard written notation. But arithmetic is another story; the extreme lack of uniqueness would have made even simple computations appear ill-defined. Perhaps it is for this reason that, in mathematics, it was used primarily to present the immutable results of calculations that could be done once and for all, probably in writing using number symbols, and then codified and committed to memory (the sine table for instance), or to express universal constants like the coefficients in the power series expansions of trigonometric functions.

Well before the advent of *kaṭapayādi*, another way of denoting numbers called *bhūta-saṃkhyā* became popular and remained in use for a long time, both in mathematical and non-mathematical contexts. What is done in this method is to denote a number by the name of an object or concept which is associated to that number in reality or by convention. Thus 'eyes', 'arms', 'twins', 'sun-and-moon' (*sūrya-candra*), etc. can be names of 2, 'sky', 'emptiness' (*śūnya*), etc. are 0, 'gods' is 33, and so on. Its origins are probably to be found in the comparison counting of the *Taittirīyasaṃhitā* in which an object is identified with the cardinality with which it occurs, 'four feet of cattle', etc. (see the next chapter). Obviously these metaphorical numbers have nothing to do with the

decimal pinciple, are extremely limited in their applicability and unsuited for serious mathematical or scientific purposes.

4.5 The Choice of a Base

The point made earlier, that the mental processes involved in apprehending and working with numbers are different for numbers less than and greater than the base, is clearly independent of the particular realisation employed. For numbers above the base we can rely on our ability to apply rules whereas such a prop is unavailable for numbers below the base; they have to be 'intuited' in one way or another. It is a matter of daily experience that the human ability to take in accurately, 'at a glance', the number of objects in a collection – to be distinguished from reading symbols representing that number, which requires learning and memory – is limited to relatively small numbers, perhaps varying slightly from individual to individual, and depending on externalities like an imposed pattern through sensory cues. If number cognition is an innate human faculty, that faculty certainly does not extend to very large numbers.

There is thus a loosely defined limit on how large a number can be that could serve as the ideal base, ideal meaning that we can trust our senses to directly and reliably apprehend all numbers below the base. Number perception is a very active field of research, but it seems that sound experimental evidence, either neuro-scientific or behavioural, that can help in identifying an approximate cognitively defined ideal base is not easy to come by;[2] among other things, the design of the experiments generally does not make allowance for the different mental mechanisms by which small and large numbers are apprehended.

In the absence of decisive scientific evidence, the best we can do then is to look around for whatever scraps of information we can gather from areas of human activity that involve counting. Within the Indian setting, one obvious place to look at is the organisation of sequences of long and short syllables that make up the rigidly codified metrical structure of Sanskrit verse. The combinatorical properties of the prosody of Sanskrit verse forms were analysed by Pingala around or just before the 2nd century BCE but quite a few of the metres he studied were already in use in the *Ṛgveda.* Apart from duration, long or short, Vedic Sanskrit assigned also a pitch to syllables. The resulting auditory cues impose a strong pattern on the chanting of the Vedas, making it quasi-musical and, thereby, a powerful aid to memory. The relevant fact for us is that the Vedic metres have, as do the majority of metres in Indian languages

---

[2]The interested reader can get an idea of the current state of research and trace the literature from, among many articles, S. Dehaene, "Origins of Mathematical Intuitions. The Case of Arithmetic", Ann. N. Y. Acad. Sci. vol. **1156** (2009), p. 232; and L. Malafouris, "Grasping the concept of number: how did the sapient mind move beyond approximation?", in I. Morley and C. Renfrew, *The Archaeology of Measurement*, Cambridge University Press (2010). One has the impression that cognitive scientists are still struggling with 'controls' , with isolating auxiliary factors which are relevant from those which are not.

to this day, a syllabic length varying from 8 to 12 for each line or foot (*pāda*). In the learning and performance of music, similarly, the rhythm cycles (*tāla*) commonly employed vary from 7 to 16 beats per cycle, grouped into two or more subcycles. It is a matter of common experience in India that a trained musician is always aware of where in the *tāla* cycle he or she is at a given moment and deviations are easily spotted, just as are metrical errors in the recitation of verse.

We can strengthen these meagre indications by recalling again that the Indus Valley culture probably used 8 as the base (with a special role for 10) and that in Old Babylonia, even though the base was an unwieldy 60, a single independent symbol for 10 made it unnecessary to take in more than 9 objects at a time. Both these are what we have referred to earlier as literal symbolic systems in which an atomic number is represented by the same number of identical tokens rather than by an abstract symbol; that makes them good candidates for testing how high the intuitive perception of (small) numbers may extend. So it is particularly interesting that the difficulties of handling a large base, 60, were mitigated by introducing an intermediate marker and that this marker was actually 10.

All of this suggests that a number around 10 is a reasonable value for the ideal base from a cognitive point of view and that it has more to do with the imperatives of visual and auditory cognition than with the accident of humans having 10 fingers – though that may have helped in converging on precisely 10. On the other hand, when it comes to abstract representations relying on learning and memory rather than intuitive perception, and in which a single symbol or name represents an atomic number essentially arbitrarily, there are different constraints to accommodate: how many names or symbols, with no other meanings attached to them and no pattern superposed, can the average human learn easily by heart in the correct sequence, to be retrieved effortlessly and infallibly at will, time after time? History and experience show that 10 meets all these demands satisfactorily.

5

Numbers in the Vedic Literature

5.1 Origins

The arithmetical parts of the *Śulbasūtra* apart, our only sources of information about Vedic numeracy are texts that are primarily poetic and/or religious-ritualistic. The references in them to numbers and basic arithmetic are, naturally, incidental but they are abundant, and rich in the knowledge that can be extracted from them. Of these texts, pride of place must go to the *Ṛgveda*, not only because it is the most ancient, but even more so because it is a veritable showcase for the methods by which the nominal representation of decimal numbers was implemented through the rules of grammar. The other important source-text is *Taittirīyasaṃhitā*, the so-called southern (or Black) recension of the *Yajurveda*, less for the details of the grammar of numbers than for a general appreciation of the state of numeric culture in Vedic and pre-Vedic, perhaps pre-decimal, times. To these have to be added a handful of later, non-*saṃhitā* texts, Brahmanas (*Braḥmaṇa*) and Upanishads (*Upaniṣad*), for the occasional allusions throwing light on some specific points. (A very accessible general account of Vedic literature – the schools, the texts and their affiliations – is available in [St-Veda]).[1] Finally and still later comes the post-Vedic material such as some of the Buddhist and Jaina writings and the epics *Mahābhārata* and *Rāmāyaṇa*. The antecedents of these later developments can be traced directly back to the Vedas and so it is natural and convenient to discuss them in the present chapter.

[1]Many of the number-related issues addressed in this chapter are treated at greater length in Bhagyashree Bavare and P. P. Divakaran, "Genesis and Early Evolution of Decimal Enumeration: Evidence from Number Names in *Ṛgveda*", Indian J. Hist. Sci. vol. **48.4** (2013), p. 535.

P. P. Divakaran, *The Mathematics of India*, Sources and Studies in the History of Mathematics and Physical Sciences, https://doi.org/10.1007/978-981-13-1774-3_5

There is one important consideration to be kept in mind when discussing the chronological evolution of the ideas that we can glean from these texts, especially the two *saṃhitā*s. The generally accepted dating of the *Taittirīyasaṃhitā* as slightly later than the *Ṛgveda* does not mean that everything in the former is later than everything in the latter. In a sense, the date of ca. 1100 BCE usually ascribed to the *Ṛgveda* is notional; it obscures the near certainty that many or most of the poems were composed in the centuries before their compilation, first into the ten individual Books and finally into the whole *saṃhitā*. Likewise for the other *saṃhitā* texts. Added to this is the tendency of archaic ideas and practices to persist in a culture long after more efficaceous replacements for them came into use. It should not therefore be surprising to find in these early texts occasional instances of less evolved methods of dealing with numbers coexisting with later, more sophisticated, developments. As we shall see, this is particularly so in the case of the *Taittirīyasaṃhitā*, which is dominantly a compendium of descriptions/prescriptions for the performance of often complex rituals, involving a great deal of counting. Vedic ritual was (and always has been) very resistant to change; perhaps for that reason, the *Taittirīyasaṃhitā* turns out to be a more fruitful source for the pre-decimal antecedents of counting than the *Ṛgveda*.

It has already been noted (Chapter 4.3) that cultures which eventually graduated to a system of full-fledged place-value counting probably passed through an earlier phase of matching individual objects of two sets one-to-one, not caring for the precise cardinality of either. In case the two sets did not match exactly, one would know which was the larger and it would have led to a determination of order, by magnitude, among numbers. The natural next step would be to ask: by how much is one set larger than the other? and so to compare the set of leftover objects with the smaller of the two sets. (A plausible scenario for how the Euclidean algorithm came into being?). From such a rudimentary and qualitative understanding of relative magnitudes to the idea of picking one convenient fixed number by which to measure all numbers larger than itself in an absolute sense was, once again, a decisive step forward, as definitive a step as that of fixing, say, the standard meter to measure all lengths with.

The *Ṛgveda* has a few instances of number-matching, the most evocative being a brief passage from Book I, poem 50: " ... you cross heaven and the vast realm of space, O sun, measuring days by nights."[2] The passage is very apt in our context in that there is no reference to the number of days or nights (that occurs in another hymn in the same Book). It is also interesting for the use of the verb 'measure' (*māna*) – the Sanskrit word for 'decimal' has always been *daśamāna*, measuring by ten. Book I is thought to be one of the two later Books and it is strikingly unexpected to find in it a figure of speech evoking an earlier, more primitive, method of direct comparison.

[2]Translation by Wendy Doniger O'Flaherty, *The Rig Veda. An Anthology* (Indian edition), Penguin Books, New Delhi (2000). Readers who wish to have an idea of the proliferation of numbers in the *Ṛgveda* without reading all of it can just sample the poems in this selection.

In some contrast to the rarity of explicit number-matching in the *Ṛgveda*, the *Taittirīyasaṃhitā* is replete with examples of sets of very diverse objects sharing a ritually ascribed unity through the mere fact of having the same cardinality. Here is a typical passage, out of very many (from Part V, Chapter 1 in the translation of Keith [TaiSam-K], slightly edited):

> He (the sacrificer) offers the Savitr (pre-dawn Sun) offerings, for instigation. He offers with an oblation ladled up four times, cattle have four feet; verily he wins cattle. The quarters are four, verily he finds support in the quarters. ... For them they kept this oblation ladled up four times. ... In that he offers what has been ladled up four times, he delights the metres and they, delighted, carry to the gods his oblation. ... This is the mastering of the sacrifice. ... These offerings to Savitr number eight; Gayatri (the metre) has eight syllables. Agni (the ritual fire) is connected to Gayatri; verily he does not abandon ... Agni as the deity. ... With four verses he takes up the spade, the metres are four; verily he takes it up with the metres. ...

I have abridged the passage with the intention of highlighting the numbers and their magical association with wholly disparate objects; it is as though the mystical power being awakened in the ritual resides in the idea of objects all sharing the same number, in the number itself. Vedic literature, early and not so early, is full of such instances, not always so explicit, of what the distinguished Indologist Jan Gonda has called "harmonies of numbers". One is tempted to say that passages such as these mark the dawning of the realisation that the faculty of being able to count leads to and derives from the abstract notion of cardinality. It was not the first time in history, nor was it to be the last, that magical powers are bestowed on new life-changing scientific and technological insights.

A second passage from the *Taittirīyasaṃhitā* is frequently quoted for its naming of the powers of 10 up to 10^{13}, the earliest record we have of powers greater than the 10^4 of the *Ṛgveda*. It occurs in Part IV, Chapter 4 (concerned with the laying of the bricks for the ritual altar), and is prefaced by the laudation (also in Keith's translation), " Thou art the measure of a thousand, thou art the image of a thousand, thou art the size of a thousand, thou art the replica of a thousand, thou art of a thousand, for a thousand thee!" Thousand here probably means innumerable, which is a usage of interest in itself as we shall see in the next section, but here we are concerned with what follows immediately:

> May these bricks, O Agni, be milch cows for me, one, and a hundred, and a thousand, and ten thousand (and so on, each power of 10 having its own individual name until *parārdha* is reached which, here, is 10^{13} if it is assumed that no intervening power of 10 has been dropped as 10 itself has been). May these bricks, O Agni, be for me milch cows, sixty, a thousand, ten thousand, unperishing; ye (the bricks) are standing on holy order, increasing holy order, ..., full of power.

Aside from the implicit recognition that numbers can be made to increase by taking higher and higher powers of 10 in (holy?) order, what is remarkable in the passage from our viewpoint is the identification of the number of desired boons (cows, the quintessential Vedic symbol of wealth), and the same number of sanctified bricks offered in exchange, with that number itself. That such a supposedly earlier stage of counting by matching occurs alongside one of the hallmarks of evolved nominal counting, the naming of high powers of 10, need not, as we have noted, be greatly surprising. As for the notion of increasing power being associated with higher and higher powers of 10, let us also observe that the name for a thousand, *sahasra*, means literally 'that which possesses power (*sahas*)'.

Before moving on to the details of the number naming paradigm, we may also wonder whether the Vedic texts explicitly acknowledge the need for a solution to the problem of the cognition of atomic numbers since that is a prerequisite for rule-based enumeration. There are in fact strong hints in the *Ṛgveda* that the question was confronted (and solved in its own fashion). The line of reasoning is somewhat indirect but, given its importance, it is worth describing even if only in outline.

The word *saṃkhyā*, for counting and hence for number, employed from Vedic times onwards, derives from prefixing *saṃ*, 'together', to the verb root *khyā*. Now *khyā* means 'to see' or 'to look at', also from the *Ṛgveda* onwards, and so *saṃkhyā* signifies 'to see together', take in at a glance to use an English idiom – an etymology that makes excellent sense in the context. Indeed there is a whole instructive array of words, already current in the Vedic literature, formed by conjoining one of several other prefixes to *khyā*, the meanings of all of which derive from the act of seeing (or making others see) and are hence linked to vision, literal or figurative (see Bavare and Divakaran, cited above, for the details; the word *vyākhyā* for an analytic commentary, a commentary which makes the reader see the subtle points, is an excellent example).

Having thus tied the notion of number to the sensory or mental faculty of seeing – no rules or algorithms here – how is one actually to perform the act of counting? The *Ṛgveda* attributes the ability to do so to a divine gift, possessed and bestowed by Agni, the god of fire and light. The passage in question is a line from Book IV, poem 2 (a hymn to Agni) in which the word *akhyad* (obviously derived from *khyā*) occurs, literally meaning 'has looked at'. The reason why it is interesting in our context is that Louis Renou (whom I have already cited in connection with fired bricks as a sacred technology), suggested in a somewhat tangential manner that in one of the *Brāhmaṇa*s (the *Śatapatha*), *akhyad* means 'has counted', a reading which was later accepted by Frits Staal as already valid in the *Ṛgveda*. With this interpretation, the line has the sense "Agni has counted (people) as an owner of cattle has counted cattle", which fits in with the poetic context in which the line occurs.

The *Ṛgveda* has numerous passages linking Agni with the power to see for himself with clarity and to enable others to see: " by his power he makes all things manifest"; he is "light for everyone to see [by]"; he has "vision that

clarifies"; he " comprehends all things minutely"; etc. etc. The play on the causal relationship between illumination and the faculty of acute vision, taken together with the etymology of the word *saṃkhyā* that we have just gone over, will then let us adopt with some confidence the reading 'has counted' for *akhyad*. In any case, the general idea of associating light and clear vision with taking in numbers at a glance is a plausible and attractive one. From an empirical standpoint, the acknowledgement it gives to the imperative of having a means of grasping small numbers, counting without counting so to say, is even more fundamental: how else to account for this faculty beyond rules except as a divine gift or, more prosaically, as a corollary of the gift of vision? Is it just a happy coincidence that, in Greek mythology, Prometheus took not only fire from the gods but also the art of numbers?

5.2 Number Names in the *Ṛgveda*

The *Ṛgveda* has close to three thousand names of numbers scattered through its ten Books comprising one thousand and twentyeight poems. A few hundred of them are names of compound numbers, i.e., those which are neither atomic nor a power of 10. The highest power of 10 occurring in the book is 10,000 and the highest number 99,000. There are variations in the names of many of the numbers, some obligaotory for reasons of grammatical correctness, and some, very likely, dictated by metrical exigencies and hence optional. Such variations are subject to two related conditions, one primarily logical and the other linguistic. The first is that, whatever the variant any compound number name appears in, there should be no uncertainty in its reading, in the exact identity of the number so named. And, to be able to assert that that is indeed the case, one must be able to isolate, on the basis of grammar alone, the rules that govern the formation of names of all compound numbers. The latter are subsumed in the grammar of Vedic Sanskrit in general as regards nominal composition. These were organised and systematised much later, in fact in Panini's *Aṣṭādhyāyī* for the first time. (There may have been earlier books on grammar, now lost.)

Thus Panini becomes our first recourse in analysing number names in the Vedas. Though he was aware of the occasional need for exceptions to his rules (excused by their occurrence in the Vedas, *chandasi*), it is amazing how well they apply, retrospectively, to the specific problem of decrypting the number names of the *Ṛgveda*. In cases where Paninian rules do not cover a particular way of formimg a number, or allow more than one reading, we have the fall-back option of going to the *padapāṭha*, the word-for-word oral rendering of the original texts (see Chapter 1.3). And, if that also fails, there is always Sayana (Sāyaṇa), the 14th century (CE) consolidator and commentator of the Vedic corpus. It may appear paradoxical that we have to depend on scholarship closer to our times than to that of the *Ṛgveda* to authenticate its interpretation, but the continuity and robustness of oral transmission are a guarantee that we will

not go far wrong. The same oral tradition is in fact our only guarantee that the text itself is authentic.

It is obviously not possible in one short section to deal comprehensively with the intricacies of Vedic grammar as it applies to the formation and analysis of number names. The aim here is only to give a broad general picture, with a few typical examples to illustrate the points that come up.[3]

There is nothing to be said about the atomic numbers since their names are completely arbitrary and grammar does not come into the picture at all. Logically, that is largely true also of the powers of 10 though, in practice, some points of general interest are worthy of comment. The names *daśa* and *śata* are the only ones for 10 and 100 and they have no other meanings prior to their use for numbers. In contrast the unique name for 1,000 (*sahasra*), as we have already seen, is also an adjective in its own right ('powerful'). Equally interestingly, so is *ayuta* for 10,000: the verb root *yu* means 'to yoke' or 'to bind', *yuta* is its past participle and the prefix *a* signifies negation. So *ayuta* has the literal meaning 'unbound'; there is in fact at least one place in the *Ṛgveda* where the word is used in that sense. The words *niyuta* and *arbuda*, which soon after (in the *Taittirīyasaṃhitā*) became the names of 10^5 and 10^7, occur also in the *Ṛgveda* but, according to Sayana, only in a non-numerical sense.

The fact that the two highest powers of 10 in the *Ṛgveda* have names conveying an idea of bigness and power carries a suggestion that the centuries during which it was composed were also a time for exploring the outer limits of numbers as they were then understood. Such a view is consistent with another special usage. There are several examples of the use of *sahasra* in which it cannot possibly have the precise connotation of 1,000, the clearest being a verse in Book X (poem 90) referring to *puruṣa* having a thousand heads *and* a thousand eyes; almost certainly, this *sahasra* stands for numerous or even innumerable. In another passage (Book IV, poem 26), *sahasra* and *ayuta* occur in conjunction and it represents not 11,000 but, according to Sayana, *aparimita saṃkhyā*, 'unbounded number'.

The rationale of the formation of decimal numbers requires us to take up next the names of multiplicative compound numbers of the form $n \times 10^m$, which are the terms in the polynomial representation, before they are finally summed up. It is more practical however to deal first with additive composition as its grammatical implementation is more transparent. Among additive compounds, the most common and presumably the earliest to have come into use are the numbers 11 to 19. There has never been any doubt that the names *ekādaśa*, *dvādaśa*, $\cdots$, *navadaśa* are the names of 11, 12, $\cdots$, 19 (no other names denote them in the *Ṛgveda*). The rule for their formation is, besides, impeccably Paninian; a *sūtra* in the *Aṣṭādhyāyī* explicitly covers this situation: in a nominal compound resulting from a wide class of *samāsa*, *dvandva* or conjunction

[3]Those who wish to pursue the analysis in detail will find a fuller account in the article of Bavare and Divakaran cited earlier.

('and') among them, *dvi* for example changes to *dvā*: *dvi* (and) *daśa* $\rightarrow$ *dvādaśa*. (An interesting sidelight is that the same name can also be derived from the rule of composition for 'two more than ten', which belongs to a subclass of the *samāsa* called *tatpuruṣa*).

Identical considerations hold for the other numbers between 10 and 20, the only variation being due to the applicable *sandhi* rules (which, to remind ourselves, are purely phonetic, having little to do with meanings), e.g., *tri* (and) *daśa* $\rightarrow$ *trayodaśa* = 13, *ṣaṭ* (and) *daśa* $\rightarrow$ *ṣoḍaśa* = 16. These names have remained the same through history, with one exception: 19, which is invariably *navadaśa* in the *Ṛgveda*, is sometimes constructed subtractively and called *ekānnaviṃśati* (1 less than 20) already in the *Taittirīyasaṃhitā* (and similarly for 29, 39, etc.), a custom which subsequently became more or less the norm. It is worth noting that this is the only departure, in modern Sanskrit-based languages, from the original strictly additive-multiplicative scheme followed in the *Ṛgveda*.

More generally, additively formed numbers of the form $10n_1 + n_0$ with $n_1 > 1$ also conform to the same principles in their nomenclature, but sometimes without completing the process of composition, i.e., without dropping connectives like *ca*, *sākam*, *na* etc. denoting conjunction. Examples are *triṃśataṃ triṃśca* (30 and 3), *nava sākaṃ navatīḥ* (9 and 90), *triśatāna ... ṣaṣṭīrna* (300 and 60) (and many others), the last being the number of days in a year. The number 99 in particular is treated in several slightly different ways: *navatīrnava, navānāṃ navatīnām, nava ... navatiṃ ca, navatiṃ ca nava, navabhiḥ ... navatī ca*, to cite a few. The important point is that there is absolutely no room for doubt that they are all additive names. This is of some advantage; since the arithmetical operation encoded in the grammatical composition can only be one of two, addition or multiplication, eliminating the additive possibility in a given name necessarily establishes it as multiplicative. It is a fact that many multiplicative compounds can be demonstrated to be such by direct grammatical analysis, but there is sometimes scope for ambiguity and, in such cases, this filter is very useful.

Multiples of powers of 10, which are the only multiples that matter in enumeration, occur freely in all the Books of the *Ṛgveda*. Of these, the most common are, once again, presumably the first to be constructed and used, namely the multiples of 10. Generally speaking, their names have a finished form in which the implementation of the nominal composition is completed, including the phonetic *sandhi* transformations. These names are standard: 20 = *viṃśati*, 30 = *triṃśat*, 40 = *catvāriṃśat*, 50 = *pañcāśat*, 60 = *ṣaṣṭi*, 70 = *saptati*, 80 = *aśīti*, 90 = *navati*. They actually appear in this order (ending in 100) in two consecutive verses of the *Ṛgveda*, in hymn 18 of the (early) Book II, the only such systematic sequence of numbers in the whole text (and hence, incidentally, the only passage of exclusively mathematical interest). There have been more than one derivation offered for these names, the most widely known being the one in Monier-Williams' Sanskrit-English dictionary. At least some of them, for instance *aśīti* from *aṣṭa* and *daśa*, are not very convincing. The

same list occurs in one of Panini's *sūtra*s and the standard commentary (*vṛtti*) in fact characterises them as grammatically anomalous. What is again beyond doubt is that their numerical meanings are unquestionable and always have been.

When it comes to multiples of 100, 1,000 and 10,000, the first thing one notices is that they do not have one-word names, in contrast to the multiples of 10. The most common method of naming such multiples is to juxtapose the names of the atomic number and the power of 10 (with a break in pronunciation); typical examples out of very many are *dve śate* (200), *trīṇi śatā* (300), *pañcā śatā* (500), *trī sahasrā* (3,000), *catuḥ sahasram* (4,000) *aṣṭā sahasrā* (8,000), etc. (It must be remembered that the endings will depend on the grammatical contexts in which they occur as determined by the gender, number and case of the nouns they qualify). The treatment of multiples of 10,000 reflects, in addition, the fact that 10,000 itself is commonly named *daśā sahasrā* or *sahasrā daśā*; they are correspondingly expressed either as an atomic multiple of 10,000 or, more frequently, as 1,000 multiplied by a multiple of 10: *catvāryayutā* = *catvāri ayutā* ($4 \times 10,000$), *sahasrā triṃśatam* ($1,000 \times 30$), *ṣaṣṭiṃ sahasrā* ($60 \times 1,000$) and even *navatirnava sahasrā* (99×1000, which happens to be the largest number in the *Ṛgveda*).

An immediate question is: how do we know that all these are multiplicative rather than additive compounds, why is *aṣṭā sahasrā* 8,000 and not 1,008 for instance? A first answer is that they cannot be additive because they do not respect any of the known rules that define addition. The conclusion can be made definitive by direct though somewhat tedious grammatical analysis. Numbers in the *Ṛgveda* generally occur as adjectives, so many of such-and-such objects, say cows, and are considered as the outcome of an action, the act of counting. Since multiplication of m by n with m as the prefactor results from the m-fold repetition of the counting of a set of n objects, in the expression '$m \times n$ cows' n continues to function as an adjective while m is to be considered as a numerical adverb qualifying n. There is an explicit Paninian rule that covers numerical adverbs denoting repetitions (*āvṛttivācaka*) and prescribes the suffixes to be attached to the name of the prefactor m. All the examples of the type $m \times 10^n$ cited above (and almost all in the whole of the *Ṛgveda*) follow these prescriptions. A point to note is that both factors in the product undergo declensions depending on the noun they, together, qualify; that is why they are generally not joined in *sandhi*. But there are also exceptions, very few, in which a *sandhi* is operative, as in *dvirdaśa* which comes from *dviḥ daśa* (10 repeated twice, 20) and *trīrekādaśām*, 3×11 similarly. Once the *sandhi* is effected, the two factors combine to become one word, exactly like *dvādaśa* and *viṃśati*. It is also to be noted that the grammatical distinction between the two factors is reflected in mathematical terminology. The first factor is called *guṇakāra*, later often abbreviated to *guṇaka* or even *guṇa*, the multiplier or active factor, and the second factor *guṇya*, that which is multiplied, the passive factor. This linguistic distinction did not come in the way of understanding that the mathematical operation of multiplication is commutative, from the earliest

times; we have already seen that 10,000 is described in the *Ṛgveda* both as 10 repeated 1,000 times and 1,000 repeated 10 times.

The purpose of this brief and inadequate excursion into grammatical detail, tedious as it may seem, is twofold: first to highlight the primordial importance of grammar for a proper understanding of the nominal realisation of decimal place-value enumeration, indeed in its very genesis, and then to show through some examples that grammatical analysis does in fact resolve possible ambiguities in the interpretation of number names. It is obviously impracticable in this book to carry out such an analysis of every potentially ambiguous number name, but the conclusion is nevertheless true: except for two or three cases where the analysis is inconclusive, the numbers occurring in the *Ṛgveda* are uniquely identified by their names.[4]

The results of grammatical analyses are also, in an ovewhelming majority of cases, supported by an authority earlier than Panini, the *padapāṭha*. The discrepancies, extremely infrequent, show themselves mainly as the incompatibility of the syntactical structure of identically constructed words or phrases with the accents assigned by the *padapāṭha* to individual syllables in them. There is also a singular (and amusing) instance in which Sayana is unable to decide (he says so) whether a certain number name stands for $20 + 100$ or 20×100. These are the exceptions mentioned at the end of the last paragraph. It is not unreasonable to attribute these rare uncertainties to the vicissitudes that long and complex texts are liable to suffer in transmission.

A final footnote. It is a fact that the numbers 60, 600 and 6,000 occur more frequently in the *Ṛgveda* than would appear to be warranted by chance alone – poem 18 of Book VII has a $6 \times 1,000$, a 60×100, a 60 and a 6 all in one stanza. Taken together with the second part of the bricks-for-cows count (60, 1,000, 10,000) in the *Taittirīyasaṃhitā* (quoted in section 1 above), is this an indication of the often suggested possibility of an earlier Babylonian influence? Against it are the following facts. i) Every single number name makes sense if and only if 10 is the base; all of the names are derived from the polynomial representation with 10 as the 'variable'. ii) In particular, the name of 60 is not in any way distinctive: it is 'six of ten', and it is followed by 'seven of ten' (not 'sixty and ten') and preceded by 'five of ten' in the list of multiples of 10 from Book II cited earlier in this section. iii) The powers of 60 have no special prominence (unlike the powers of 10); 60 itself takes its appointed place among the multiples of 10, $60^2 = 3,600$ does not occur at all and 60^3 is bigger than the biggest number in the *Ṛgveda*. If there was an early Babylonian influence on Vedic numeric culture, it is not visible in the number system that we see displayed in such prolific detail in the *Ṛgveda*.

---

[4] A historically interesting example of the hazards of too facile a reading is the case of Colebrooke (in his book of selected chapters from Brahmagupta and Bhaskara II, [Co]) and Burgess (in his critical translation of the *Sūryasiddhānta*, [SuSi-B]) who both misread *aṣṭaśata*, referring to the number of verses in *Āryabhaṭīya*, as 800. See Chaper 1.4 for the background. If 800 was meant the name would have been *aṣṭā śatā* with a break, as in *dve śate* for 200, *pañcā śatā* for 500, etc., see above.

To summarise then, there can be no doubt at all that, by the time of the composition of the earliest individual poems of the *Ṛgveda*, the Vedic people had achieved a clear understanding of the principles underpinning decimal enumeration and its practical realisation through a rule-based system of number names. These linguistic rules are as sharply defined and their application is as precise as the arithmetical rules they encode. Though no one wrote a treatise on what we may call the grammar of numbers as Panini did for the grammar of Sanskrit, the evidence of the *Ṛgveda* would seem to suggest that the two grammars had a mutually influential symbiotic relationship in their formative period. The system of numbers that emerged out of this intimate interaction and its epistemic basis did not change much over the millennia until, at the very end, the demands of the new mathematics of the Nila school forced a rethinking of the foundations (Chaper 4.2).

5.3 Infinity and Zero

We have already noted that the highest power of 10 in the *Ṛgveda*, 10^4, was rapidly overtaken by the 10^{13} of the *Taittirīyasaṃhitā*. The *Taittirīyasaṃhitā* has, in fact, another power list in its Part VII, Chapter 2, longer than the earlier one and full of interest of its own. It is a count of oblations (to Prajāpati) accompanied by a laudation (*svāhā*) or, in Keith's reading, laudations of the numbers themselves, beginning with 100 ("To a hundred hail!", etc.). The names are the same as in Part IV until 10^{12}, but 10^{13} is not *parārdha* but *uṣas* ('dawn') and the subsequent names also all have literal meanings (as indeed do most of the names already in the earlier list). Assuming again that no powers are skipped, it ends with 10^{19} which is named *loka* ('world'), followed by a universal laudation *sarvasmai svāhā*, "All hail!" in Keith's translation.[5]

From the point of view of the history of numbers, two things stand out about this list. The first is that, unlike in the list of Part IV, there is no sign of number-matching in the way it is put together; in Keith's translation, it goes: To a hundred hail! To a thousand hail! To ten thousand hail! and so on. Much more intriguingly, the geometric progression of the powers of 10 comes at the end of a succession of other sequences which are (partly) in arithmetical progression and so is of interest purely mathematically – i.e., quite aside from whatever cultural historians might make of the interplay between ritual and mathematics. Postponing for the moment a more detailed look at the arithmetical sequences and what they can tell us, let us first trace in outline the evolution of the power lists in the literature that came after the *Taittirīyasaṃhitā*. There are quite a large number of texts, some religious and some more or less secular, that carry such lists, the numbers not always identified by name nor always progressing in multiples of 10. A general trend is that with the passage of time

[5]Keith does not consider that *uṣas*, ..., *loka* are numbers but takes them at their literal meanings, thus ending the sequence at 10^{12}. In the context of the passage and of the names of powers of 10 in general, it is difficult to see why.

the tabulated sequences extended to ever higher powers of 10 until about the 2nd or 3rd century CE, after which time the fascination with bigness wore off. Mathematicians (Aryabhata, Bhaskara II and Jyeshthadeva, to name only the most prominent) continued to cite fairly long (but shorter than the second list of *Taittirīyasaṃhitā*) tables, either to underline the theoretical basis of decimal enumeration or because they needed them in their work.[6]

These power lists occur in many different contexts. Among the earliest in the period following the *saṃhitā* texts is a passage from the *Taittirīya Upaniṣad* that puts numbers to the bliss of attaining *brahman* (a word that can be rendered variously and always inadequately, here perhaps best as the transcendental and universal godhead of late Vedic philosophy). It begins by defining the happiness of a young and vigorous man, endowed and accomplished in every way, as one unit. One hundred units of this perfect human happiness is then stated to be the happiness of a celestial being, and so on, until, in ten steps of hundred-fold increase, one arrives at the measure of the bliss of *brahman*, equal to 10^{20} units of the bliss worldly humans are granted. There are no names given to the numbers except for 100.

It is difficult to give dates, even approximate ones, to most of the other relevant texts that were composed in the seven or eight centuries following the early Upanishadic period. The main reason for this is the high probability that, in the form in which they have come down to us, they are the end result of many accretions and emendations, over a substantial period, to an original core. Only a few typical examples are mentioned below, more to give an idea of the general preoccupation with enormously large numbers and their potential infinitude during this entire period than to establish a pattern of evolution. Among the smaller of these truly large numbers is 10^{28} which, according to the Jaina canonical work *Anuyogadvāra Sūtra*, was the total number of people in the world. The *Rāmāyaṇa* in contrast thought really big; it has a long passage on the strength of Rama's army before the invasion of Lanka.[7] The count begins at 10^5 (named a hundred thousands) and proceeds somewhat erratically in steps of varying powers of 10, once by 100, once by 1,000, otherwise by 10^5, to arrive at the final figure of 10^{60} (in my count) which is given the name *mahaugha* (all of the powers that come up at the intermediate stages of the count also have their own proper names). The *Mahābhārata* too flaunts large numbers in more than one place, the number of trees in a forest (in the story of Nala and Damayanti) for example.

Large as the numbers in these stories may be, it is to the Mahayana (Mahāyāna, the reformist subsect that split from the parent in the 3rd century

[6]Several such tables, covering the entire span of Indian mathematics, are given in Hayashi's edition of the Bakhshali manuscript ([BM-H]). Some lists from Buddhist and Jain literature can be found in R. C. Gupta, "Wolrd's Longest Lists of Decuple Terms", Gaṇita Bhāratī **23** (2001), p.83. There does not seem to exist a comprehensive census of large powers of 10 nor is there always unanimity about which name denotes which number.

[7]The passage is quoted in Sanskrit in S. A. S. Sarma, "Vedic Numeral Sytem including *Śūnya*", in [Sunya]. Some of the other lists I cite in this section are also given in this article.

BCE) Buddhist literature that we have to turn for the most elaborately contrived ways of coming to terms with a fact that was increasingly becoming evident to learned men: that there really was no end to the numbers that can be imagined; only, their concrete manifestation was limited by the finite human capacity to talk about them and to find names for them. The Mahayana world-view embraced a cosmogony in which the world of the here and now is just one in a multiplicity of worlds without end, temporally or spatially. Such a picture, perhaps, both derived its inspiration from and fed into the notions regarding the endlessness of numbers which these stories are struggling to comprehend; there is an almost palpable sense of wonder in many of them.

The most dramatic, in content and narrative style alike, is a story from *Lalitavistāra*, a mythical-miraculous account of the life of the Buddha, dated to the 3rd century CE in its final form but with many signs of its having begun existence as a much earlier work. The particular episode relates to a contest the prince Siddhartha (Siddhārtha), the future Buddha, is obliged to engage in in his youth (as a test before his marriage). The story relating to numbers is that he is examined by the greatest mathematician in his father's kingdom, by name Arjuna, about his knowledge of systems of enumeration. In effect, Siddhartha's elucidation of what was clearly thought to be a very arcane science amounts to getting first to a really large number multiplicatively in steps of 100 (10^{53} from the description in the story) and then using that as the unit for further multiplicative increases, the whole process to be iterated 10 times. And it does not end there. There is obviously not much point in trying to determine the largest number the enumeration might lead to; that was probably not the purpose. The story has many graphic details of physical objects – grains of sand in river beds (and in the beds of as many rivers as there are grains of sand in one), atoms in the kingdom of Magadha (and in all kingdoms put together), etc. and, at the other extreme of abstraction, the totality of all the worlds of the Mahayana cosmogony – whose numerical magnitudes can only be grasped by a being on the way to enlightenment. It is best just to say that the idea being conveyed is that there are numbers that are practically beyond count; the word *asaṃkhyeya*, 'uncountable', actually occurs in the narrative in a context that merges the very small with the very large, in speculations about the total number of atoms in all the worlds.

These intimations of the infinite from the non-mathematical literature are especially valuable because they span a period, from the *Śulbasūtra* onwards, that produced hardly any mathematical text properly speaking. The surprise is that the non-scientific writings of the time are so concerned with and so knowledgeable about issues regarding how the simple act of counting leads directly to questions of such mathematical – Arjuna is described as the foremost mathematician, *gaṇaka*, of the kingdom – and metaphysical import. For the Buddha, to acquire learning was the first step on the path to enlightement, just as those who put together the *Taittirīyasaṃhitā* were learned men before they were sages and seers. To know numbers was part of being learned.

Viewed in this light and stripped of metaphysical and ritual overtones, it is plausible to see in the two power lists of the *Taittirīyasaṃhitā*, especially the second one, the first awareness of the endless progression of numbers. As noted earlier, it occurs in Part VII, Chapter 2, following immediately after several other sequences. The first of them starts as a count of the natural numbers and goes (in Keith's translation)[8]: "To one hail! To two hail! ... "(until 19) and continues "(To nineteen hail!), To tweentynine hail! ... " (until 99), ending with "To a hundred hail! To two hundred hail! All hail! (*sarvasmai svāhā*)". (The dots are inserted by me to save writing; the numbers they stand for are present in the original). The sequences of numbers hailed in the following passages are (1, 3, 5, ..., 19, 29, ..., 99, 100) (and another identical one but dropping 1 at the start); (2, 4, ..., 20, 98, 100); (4, 8, ..., 20, 96, 100); (5, 10, 15, 20, 95, 100); (10, 20, ..., 90, 100): (20, 40, 60, 80, 100); and (50, 100, 200, ..., 900, 1,000). It is after these nine largely arithmetical sequences (which they are apparently meant to be, once we fill in the gaps) that we get the power sequence (100, 1,000, ..., 10^{19} = *loka*, the world itself). Every single one of them ends with the coda *sarvasmai svāhā*.[9]

The first reaction one has on reading these incantations is to see (or, rather, hear) in them a numerical equivalent of the play with words and syllables of the *padapāṭha* and the *prātiśākhya*. There is a sense of exploration: how can numbers be played with so as to make them reveal their structure better? And the answer seems to have been that, however they are manipulated, their is no finality to them; the abrupt closure of the chants with *sarvasmai svāhā* is like an acknowledgement of that inescapable truth, whichever of its possible meanings one gives to *sarva*.

The preoccupation with the very large (*asaṃkhyeya*) of the early centuries CE soon led to a sharper and more precisely mathematical articulation of the idea of infinity by the 6th century CE. The clearest evidence comes from the investigation of the solutions in integers of the indeterminate equations of the 1st and 2nd degree introduced by Aryabhata and Brahmagupta. That they had an infinite number of solutions was quickly recognised and authors in the ensuing period have frequent references to their being endless (*ananta*). It comes as something of a surprise therefore that the first explicit statement to the effect that decimal numbers themselves, in the abstract, are without end comes only in the last phase of Indian mathematics (to my knowledge). As we will see in Part III of this book, that fact is of crucial importance in the work of the Nila school on calculus: the infinitesimally small was defined by dividing a finite (geometric) quantity, say the length of an arc, by a large number and letting the divisor increase without bound. After listing the names of the first 17 powers

[8]The original versions of most of the lists are given in the article of S. A. S. Sarma cited earlier.

[9]Keith's translation is perhaps too free; if *svāhā* is given its original Vedic meaning of oblation, the priest is not hailing the numbers themselves in these passages, but counting the oblations as he makes them. That would strengthen the case for a strictly numerical interpretation (though, naturally, indefinite) of *sarva*.

of 10, *Yuktibhāṣā* goes on to say: "Thus if we endow [atomic] numbers with multiplication [by powers of 10] and positional variation, there is no end to the names of numbers and hence we cannot know all the numbers and their order". It is striking that the approach to infinity is multiplicative rather than through iterated succession, a notion that *Yuktibhāṣā* was familiar with as we saw in Chapter 4.2.

It is something of an anomaly that a great deal more has been written about the Indian idea of 'nothing' (which has many names in Sanskrit, not all of them necessarily perfectly synonymous) and its mathematical expression, the zero (whether nominal or symbolic), than about infinity. The subject also seems to polarise opinion, to a greater degree than usual, about the original source of the notion of zero as a number – indigenous? if not, from where? and when? Part of the reason may be that it is of course difficult to find solid – read written – evidence pertaining to a period which has left no durable written records of any sort and in which writing was probably not a skill practised by the learned. More fundamentally, in an orally literate culture, the idea of the mathematical zero is irrelevant in enumeration (as I have been at pains to emphasise) and of secondary importance in computation. The number names of the *Ṛgveda* have no need or place for a zero. Nor is it required in arithmetic with positive integral (or even rational) numbers: the result of multiplying, for example, (m_2 hundreds and m_0 ones) by (n_2 hundreds and n_0 ones) is just (m_2n_2 ten thousands and $m_2n_0 + m_0n_2$ hundreds and m_0n_0 ones), reducible to the standard form by means of the carry over rules, with not a zero in sight. This may appear to be a trivial point but it is one of the differentiators between the oral and written representations of place-value numbers; used as we are today to an overwhelmingly symbolic approach to numbers and their manipulation, it takes a conscious effort to keep this distinction in mind. Historians, in particular those who anchor themselves in the written tradition of Mesopotamia, do not always take account of the fact that the functional roles of the zero in written and oral cultures are very differently manifested.

A second point is the following. It is often said that the zero has two different roles in our understanding of numbers: as a natural number, the counting number preceding 1 (which is essentially the same as the cardinality of the empty set) and as one of the possible entries in any place in place-value notation, to indicate that that particular place is empty. The distinction is more a matter of convention than of logic. The notation $n_kn_{k-1}\cdots n_1n_0$ is no more than an economical short-hand for the polynomial expression $\sum_{i=1}^{k} n_i 10^i$ where each numerical coefficient n_i can range over $0, \cdots 9$. Mathematically they are strictly equivalent; an empty place just corresponds to a term that is absent in the polynomial because it (its coefficient) is equal to the number 0. If the scribes of the past had demarcated the succession of places in a less confusing way than the Old Babylonian gaps, say by making compartments linearly arranged as the Indians did from at least the time of the Bakhshali manuscript, they could very well have done (and sometimes did) without a written zero; the symbol for zero would just have been the absence of a symbol. Such a notation

would of course be clumsier than the current universal practice. But the fact remains that the appearance of a written symbol for zero is contingent upon the conventions of a particular way of representing the idea of 'nothing'. It is not the signal of a conceptual breakthrough; the conceptual advance was the polynomial representation and that happened much earlier.

The real question therefore should be: what historical evidence do we have for the realisation that zero (0) is to be treated like any other number (having certain special properties on account of its being the neutral or identity element in addition), without the metaphysical overtones that the idea of the void seems always to have carried in India, and when did that realisation dawn? The *Ṛgveda* definitely has no name which unambiguously represents this mathematical zero. Strictly speaking, it did not have to, as we have seen, but given how clear its understanding of the structure of decimal numbers was, it is still something of a surprise. Nor does the *Taittirīyasaṃhitā*. When, later, it did get a name, it got not one but many, in strong contrast to the unique names of all the atomic numbers from the *Ṛgveda* onwards. And all the names of zero had other prior meanings, generally connoting absence or emptiness (some going back to the early Vedic texts): *śūnya, kha*, *ākāśa*, *bindu*, etc. etc., and even *pūrṇa*, which otherwise means 'full' or 'complete', the exact opposite of 'empty'. It is difficult to escape the feeling that they are more in the nature of metaphoric numbers (*bhūtasaṃkhyā*) than proper number names like *eka* for 1; that, mathematically, the zero was an afterthought.

The crystallisation of the notion of emptiness into a concept having a precise numerical significance – a number like any other, more or less – must have been a gradual process. As far as we can tell from the texts, a well-defined notion of an absence or vacancy first made its appearance in grammar (probably following from certain rules of omission in the chanting of Vedic *mantra*s and in the recitations of the *padapāṭha*), in fact in Panini in a celebrated *sūtra*, *adarśanaṃ lopaḥ*, "*lopa* is that which does not appear". The rule stipulates the circumstances in which an infix that otherwise should be present disappears (becomes *lopa*) from a phrase but it does not leave behind a gap. There are other grammatical situations in which *lopa* is used to denote omission, for instance a species of nominal composition called *madhyama-pada-lopa samāsa* (the *samāsa* with the middle word missing), in which the middle connective is dropped when compounding three words, again without leaving a tell-tale gap. In this sense, the *lopa* is different from the zero in enumeration; the zero was either never present to begin with as it was not needed (in the oral representation) or its presence has to be indicated, even if by the absence of a symbol (in a visual, including written, representation).[10]

[10] Much has been written about whether Panini's *lopa* directly prefigures the mathematical zero, surely inspired in part by his formal ('set-theoretic') approach to grammar; modern grammarians in particular sometimes say that, where the *lopa* rule applies, something is 'replaced by zero'. Nothing actually takes the place of something that pre-existed, it just disappears. Two recent protagonists of the view that *lopa* is the ancestor of the zero are Frits Staal, "On the Origins of Zero" in [SHIM] and M. D. Pandit, "Reflections on Pāṇinian Zero"

The next landmark in the evolution of the zero is the use in Pingala's *Chandaḥsūtra* of the word *śūnya* in a context which leaves little room for doubt that what was meant is the number 0. Reasonable datings of the *Chandaḥsūtra* based on the style and structure of the text places it in the 2nd - 4th century BCE, i.e., within two or three centuries after Panini. The work is concerned, as mentioned already, with a mathematical system of classification of Sanskrit metres. Postponing a look at the combinatorial aspects of its content to section 5 below, let us focus here just on the use of *śūnya.* In describing a procedure for computing the total number of 'words' of a fixed length that can be formed from an 'alphabet' of two 'letters', Pingala is required to distinguish between even and odd numbers that arise in the intermediate steps. His prescription is: tag the even numbers so arising by the label *dvi* and the odd ones by the label *śūnya.* The final formula for the number of words depends on this separation into two subsets, but not on the literal meanings of *dvi* and *śūnya* – they might just as logically have been denoted by abstract markers à la Panini. For us the significance of the *sūtra* lies solely in the pairing of the two tags, *dvi* and *śūnya.* The only meaning of *dvi* being the number 2, one will have to be perverse to suggest that *śūnya* is anything but a number; and that number can only be 0. Pingala is thus the first on record to give *śūnya* the mathematical connotation of 0, which thence became its proper name as a number.

It is necessary to say here that there is nothing in the *Chandaḥsūtra* itself to suggest that zero, or for that matter two, was written down as a symbol, though Pingala lived at a time when the Brahmi script almost certainly was in existence – the Ashokan edicts were only a century or so into the future, perhaps even contemporaneous. For the first persuasive though still indirect reference to the written zero, we have to wait until about the 4th century CE. It comes in the form of an isolated simile in a romantic work of fiction, *Vāsavadattā* of Subandhu, comparing stars in the night sky to *śūnyabindu.* Literally, a *bindu* is a dot before it becomes the metaphoric number 0. The composite word *śūnyabindu* makes sense only if we analyse it as 'the dot that is the zero'. Taken with the fact that *bindu* is a somewhat unusual choice as a metaphor for zero (most of the others have a direct connection to emptiness), the implication is very strong that the dot refers to a written zero, exactly as it is depicted in the Bakhshali manuscript. This will not of course mean, in itself, that Bakhshali is to be dated in the 4th century CE but lets us place an upper limit on the date of introduction of a symbolic zero at around 400 CE. But it cannot be much earlier; the reason has to do with Brahmi numbers and the principles governing their forms – a subject which seems not to have

in [Sunya]. Pandit suggests that *lopa* is specifically the exact grammatical equivalent of 0 as a place-entry. It is obvious that such an equivalence cannot go much beyond the observation that some numbers have zeros in their place-value representation just as some phrases result from dropping an infix; it cannot be extended to the zero in its arithmetical meaning, and so is of limited value. In any case, we have seen that mathematically there is no real difference between the two roles of zero.

received much attention in the literature – and throws a good deal of light on the evolution of a truly positional number notation.

The early Brahmi inscriptions (up to about the 3rd century CE) with numbers in them – and there are lots of them – do not employ a zero symbol. What they have instead are individual *single* symbols for the powers of 10 (in addition to the atomic numbers), combined in a variety of ways, depending on the epoch and the region, to write composite numbers. Most systematically treated are the multiples of 100 and 1,000, which are written as symbols for 100 and 1,000 modified by attaching 'suffixes' (see the illustrations in the book of Datta and Singh [DS]). They are thus nothing but a faithful transcription of number *names*, constructed grammatically, into visual symbols in which the rules of multiplicative composition are implemented by suffixing (additive composition is just juxtaposition) – a hybrid representation twice removed one might say, from concept to name to symbol (for the name, not for the concept). That, as we have seen, does not need a zero. It also means that the positional notation, not just the symbolic zero, was not current at the time of these inscriptions.

All in all, we can conclude with a degree of confidence that the concept of zero as a number was grasped some time before 300 BCE (but probably after 500 BCE) and that its first written manifestation came about in the century preceding 400 CE.

Keeping aside the enigmatic Bakhshali manuscript, the first dated representations of an inscribed 0 that we have are not from geographical India but in inscriptions on rock from places in modern Cambodia and Indonesia, in the form of dates. The dates span the years 683 to 687 CE and are given as years in the (mainland Indian) Śaka calendar (605, 606 and 608, all, fortuitously, including a 0), whose zero year begins in 78 CE. The scripts used are local but the language of the inscriptions is Sanskrit. Since large parts of southeast Asia had become Indianised in their culture, language and religion well before the 7th century, neither the use of Sanskrit nor of the Śaka calendar should be a surprise; there is no doubt that the zero of these inscriptions is the Indian zero.[11] Interestingly, it is represented both as a dot (an incised indentation in the stone) and as a small circle; so, the transition from the *bindu* to our familiar 0 was well underway by the 7th century. Why the circle eventually displaced the dot must remain a matter for speculation, perhaps because of the difficulty of incising a point with the tip of an iron stylus on palm leaf which became the preferred method of writing in the plains of India. In any case the circular 0 does not appear in India proper until the well-known stone inscription dated 876 CE from Gwalior in which 0 occurs twice both, happily, at the end

[11] These facts were brought to light in 1931 by Georges Coedès, an authority on the cultural Indianisation of what used to be called Indochina, and subsequently taken note of by many writers. Neverthess, Joseph Needham's celebrated many-volume study *Science and Civilisation in China*, says (in 1959) at one point that the finding of the first dated zero "on the borderline of Indian and Chinese culture-areas can hardly be a coincidence" – the conclusion is left for us to draw. Interested readers will find more detail in Frits Staal, cited above.

of numbers (in the place of ones). Neither number is a date, contrary to the general rule in royal inscriptions; indeed, the relative rarity of the written zero in inscriptions may have something to do with the fact that only about 20 per cent of three-digit (year) numbers can have a 0 in them. It is also in the period preceding the Gwalior inscription that decimal enumeration with 0 began to be propagated outside cultural India, to Persia first and then, ironically, to Mesopotamia, the birthplace of place-value numbers.

5.4 Early Arithmetic

There is very little to say about arithmetic in the *saṃhitā* texts because there is very little of it, apart from the evidence they provide of the arithmetical insights on which decimal number nomenclature is based: multiplication of powers of 10 by atomic numbers and the addition of such multiples. These are easily performed and do not involve any (rule-based) algorithm such as long addition and long multiplication.

The unique (to my knowledge) example in the *Ṛgveda* of general addition, i.e., not of powers of 10, is $35 = 15 + 20$, (it is not given as $3 \times 10 + 5$) from Book X (addition is indicated by the conjunction *sākam*). There are quite a few examples of general binary multiplication and they are all elementary, 2×5, 3×7 (often), 3×11 (once), etc. and we know that they are multiplications because they obey the appropriate rules of grammar described in section 2 of this chapter. An example, again unique, of the multiplication of three numbers $(3 \times 7 \times 70)$ occurs in Book VIII (which is rich in numbers) and it is a text-book illustration of the rules regarding numerical adverbs for counting repetitions, also discussed in section 2. To these one should perhaps add the verses from Book I which give the number of days in the year as 360 and the number of days and nights as 720, whether it was obtained as $360 + 360$ or 360×2.

The first instance we have of the operation of division is from the *Śatapatha Brāhmaṇa*, probably slightly before Baudhayana's *Śulbasūtra*, and it is of exceptional value for an understanding of early arithmetic. The passage in question is long and very graphic, ostensibly about the division of the body of Prajāpati by himself.[12] But, metaphorically, Prajāpati is also the year with its 720 days and nights at one level and, at another, the ritual altar having the same number of bricks, very reminiscent of the magical power of numbers in the abstract as invoked in the *Taittirīyasaṃhitā*. Setting aside the ritual connotations, what the passage does is to go through all the numbers from 2 to 24, checking to see if they are factors of 720, identifying those which are not (e.g., "he did not (could not?) divide sevenfold" and so on). The numbers which do divide 720 without a remainder are characterised as failures ("he did not succeed") until he comes to 24 and "there he stopped, at the fifteenth". 15 is the total number

[12] For a translation of (most of) the passage see Kim Plofker's book [Pl]. The phrases below within quotes are in Plofker's translation.

of distinct factors (including 1) and he stopped at 24, the "fifteenth arrangement [of bricks]", because he succeeded. What he succeeded in was finding the highest factor of 720 less than its square root, in other words (knowing also the quotient which is given explicitly for some of the divisions) all the factors.

Clearly, this passage could not have been written without long division, i.e., an algorithm for division of numbers larger than the square of the base (by numbers larger than the base). That the search terminates at 24 also shows that the commutativity of multiplication was understood. At a still more basic level, this is the first evidence we have for an appreciation of the process of division with and without remainder, the underlying arithmetical principle of place-value enumeration.

A direct interpretation of this exercise in division, as the text itself makes clear, is as the determination of the number of bricks in one course in an altar made of identical courses, with 30 as the minimum: "He made himself twentyfour bodies of thirty bricks each". But the astronomical interpretation is more interesting. The number of factors of 720 that are less than its square root, 15, is identified with the number of days in a *pakṣa* (the waxing phase or the waning phase of the moon) and the highest such factor, 24, with the number of half months in the year, each full month consisting of 30 days and 30 nights or 60 days-and-nights. It is curious that 15 is arrived at as the number of factorisations of 720 which by chance also happens to be one of its factors – another magical numerical coincidence? Perhaps these interrelated numbers played a part in fixing 360, already in the *Ṛgveda*, as the number of days in a year. The case can also be made that such astronomical coincidences are responsible for the relative prominence of 60 and its multiples (rather than its powers) in the *Ṛgveda*, and not directly the Babylonian base 60 (see section 2 above) which, according to modern scholarship, arose as a convenient multiple of the diverse units employed in trade.

There is a good case to be made that isolated snippets like this, taken from the non-mathematical settings in which they occur, are not a true reflection of the level of arithmetical knowledge of the times. The *Śulbasūtra* of Baudhayana and Apastamba, written not long after the *Śatapatha Brāhmaṇa*, are proof that these geometers were the inheritors of a strong arithmetical tradition as well, as we saw in Chapter 2.5. The degree of numerical skill that we may attribute to them depends to some extent on our understanding of the unstated rationales they had for their results such as the approximate rational formulae for π and $\sqrt{2}$. In particular, if some of the modern reconstructions of their thinking are close to being right, their technical virtuosity was of a high order, not to mention the conceptual advance involved in transforming the geometric problem of determining the diagonal of a square into its arithmetical equivalent of computing square roots. There is not much to add to the discussion in Chapter 2.5 of these accomplishments, except to note a few general points of historical interest.

Firstly, it is difficult to overlook the fact that both π and $\sqrt{2}$ are expressed as algebraic sums of terms, all with 1 as the numerator, and that the

denominator of a term is obtained from the preceding one by multiplication by one number. Possibly, this is an indication that each term after the first was computed as a correction to the sum of the preceding terms: there is no attempt to express the whole sum as a 'vulgar' fraction, (numerator)/(denominator). Very surprisingly, nor is it given as a decimal fraction, a sum of terms with powers of 10 as denominators, the decimal equivalent of Babylonian sexagesimal fractions. Indeed, Indian mathematics never adopted the decimal representation of fractions throughout the course of its history. Such an absence is obviously a factor to be taken account of in discussions of possible Babylonian influences – especially given the common allegiance to the place-value principle – just as the absence of parallelism as a driving force in India weakens the case for a common origin for Indian and Greek geometry (Chapter 3.4).

There is yet another significant absence. The numbers π and $\sqrt{2}$ got special attention in the *Śulbasūtra* not because of their arithmetical significance as irrational numbers but for their geometric relevance. That is natural but the Greeks who also started from the geometry went on to recognise their incommensurability property and in fact proved it for $\sqrt{2}$ (though there have been suggestions that the famous proof in Euclid's Book X is a later interpolation) by a purely arithmetical method. Despite the qualification "with remainder" attached to the rational approximations to these numbers (several such in the case of π) in the *Śulbasūtra*, there is no evidence at all that their irrationality was suspected or even that the notion was grasped – to find the first mention of irrationality we have to wait till the beginning of the 16th century (Nilakantha). Why this is so is something of a mystery, especially in view of the several methods described in the texts for finding square roots, exact in the case of squares of rationals, approximate otherwise.

At a similar level of generality, we may also wonder whether there is any indication of negative numbers in the *Śulbasūtra*. In a mathematical culture dominated by the geometry of areas, the extension of positive integers and fractions to include their negative counterparts is a less compelling thing to have done than the extension of (positive) rationals to (positive) reals. To a certain extent, our constant familiarity with negatives, as a concept and in the arithmetical operations of daily life, can blind us to the sophistication of thinking required to give them the same status as the positives. Operationally, negatives tended to be introduced in a culture when quantities were dealt with which had an obvious 'opposite'. Different civilisations adopted them (or got reconciled to them) at different times and cultural predispositions seem to have played as big a role in this process as perceived need. From a theoretical angle too, to go from $a-b$ for $a > b > 0$ to $b-a$ and further to define, say, $(-a)(-b)$ in a logically coherent way requires a generalisation of the principles of arithmetic with positives in a way that is not immediately (intuitively) obvious.[13]

[13]These issues are described and analysed in David Mumford, "What is so Baffling About Negative Numbers? – a Cross-Cultural Comparison" in [SHIM].

The *Śulbasūtra* of course has subtractions, but no negative numbers *per se*. The subtractive terms in the formulae for π and $\sqrt{2}$ for instance come after positive terms which are larger in magnitude; if the operations are performed in the order in which they are described, negative numbers do not appear even as an intermediate result. At an operational level, it would not have been obvious in Vedic times what the opposite of the circumference-diameter ratio is (directed arcs and signed angles lay in the distant future) and certainly not natural to think of 2 as $(-\sqrt{2})^2$. And conceptually as well: to go from a positive m to $-m$ by removing 1 repeatedly one must pass through 0 and there was of course no mathematical zero at the time.

The Bakhshali manuscript has negative numbers as well as the *bindu* zero – the commercial orientation of many of the problems in it fits in perfectly with the trader's need to know when he is going to break even or make a profit or a loss; all the more reason to regret our ignorance about its date. But if Bakhshali is not really early, the first explicit mentions of negative numbers in an arithmetical context (in astronomical calculations) should be considered to be in the *Vāsiṣṭha Siddhānta* as selectively presented in Varahamihira's *Pañcasiddhāntikā*. The terminolgy for positive and negative numbers, *dhana* and *ṛṇa*, is the same as in Brahmagupta (see below) and all later authors. The work itself dates from the first half of the 6th century CE but it is declared to be a compendium of five ancient *siddhānta*s; parts of *Vāsiṣṭha* are, from several indications, among the earliest portions, possibly as early as the 2nd century CE. (Chapter 6.2 has a discussion of these texts).

But whatever their origins, the true home of negative numbers (and, partly, also of 0) is to be sought in the abstract realm of algebra rather than in arithmetic narrowly defined. In a problem in which the answer is obtained by setting up and solving an algebraic equation for a quantity which is *a priori* unknown, any algebraic expression arising in the course of the calculation, even one that is linear in the unknown or variable, cannot be guaranteed to be positive for all values of the variable. The formal reason for this, as we can see today, is that the domain of numbers (integers, rationals, reals) in which such a linear problem is set up and solved has to have the structure of an Abelian group (see Mumford, cited above) with (algebraic) addition as the group operation and, therefore, an additive inverse for every number and an additive identity, 0, that bridges the gap between 1 and -1.

Viewed in this light, it is entirely natural that the first clear enunciation of the rules of signs, for dealing with positive and negative quantities together, came from Brahmagupta (in 628 CE, *Brāhmasphuṭasiddhānta*, Chapter XVIII), in a context which covers both arithmetic proper and algebra. Significantly, these rules are stated simultaneously with the corresponding rules for 0 and they are unambiguously correct and unexceptionably modern in word and spirit (including in particular $(-a)^2 = a^2$) except for the rule for division by 0: $0/0$ is stated to be 0 but, for a non-zero a, $a/0$ is to be left as it is – because it is not defined as we would say today? Even more to the point, Brahmagupta was the true pioneer of symbolic algebra, the first to set up a systematic framework

in which to pose problems requiring the performance of algebraic operations on unknowns, as well as the first to denote multiple unknowns by the syllables of Sanskrit (as distinct from their use as metalinguistic labels to designate sets *à la* Panini and Pingala and going beyond the dot and the *yāvat tāvat* of the Bakhshali manuscript). It is also not a coincidence that *Yuktibhāṣā*, in describing the next major algebraic advance (900 years later), that of the abstract definition of polynomials in one variable, prefaces its account by quoting Brahmagupta on the multiplication of signs.

The algebraic aspects of Brahmagupta's thinking cannot have been *svāyaṃbhuva*,'self-existent', discovered in a sudden act of enlightenment. It would be pleasing to know that the centuries over which the idea of negatives matured and became part of the mainstream paralleled the evolution of the idea of the mathematical zero. In the case of the zero we could reconstruct, partially but plausibly, its trajectory from Panini onwards through allusions mainly in non-mathematical works (Pingala excepted). There is no such help available so far for negatives, perhaps not surprisingly: one non-mathematical discipline where we might have expected to find a formal characterisation of the notion of opposition is logic but, during the relevant time (and later), influential schools of Indian logic, especially the *Nyāya* school which appears to have had some influence on the doing of mathematics, did not accept mutual binary opposition between a proposition and its negation, bivalence in short, as a valid principle of reasoning. Instead, especially among Buddhist philosophers (going back to the first of them, the Buddha himself), the favoured paradigm (there were other variants) was to consider, along with a proposition 'p', the propositions 'not-p', 'p and not-p' and 'neither p nor not-p' as possible alternatives. This philosophical position had other and seriously detrimental mathematical consequences in the matter of what was to be considered an acceptable demonstration, which will occupy us at some length in Part IV (Chapters 15.1 and 15.2).

A final comment about division by 0. It cannot possibly have escaped Brahmagupta that $a/b = c$ implies $a/c = b$ (it is just the rule of three) and that as b became larger and larger, c would become smaller and smaller and *vice versa*. Why then did he (and virtually everyone who followed him) not say at least that $a/0$ for $a \neq 0$ is a very (unboundedly) large number even if they shied away from thinking of 'infinity', which they knew to be unattainable, as a number? With Bhaskara II we can glimpse the beginning of some clarity: he has a name, *khahāra* ('divided by 0'), for the result of division by 0 and says in *Bījagaṇita* that its "nature" does not change when numbers are added to or subtracted from it (nothing is said about its multiplicative properties, in particular about division by *khahāra*. But in *Līlāvatī* he also says, echoing Brahmagupta, that if a number is to be divided by 0, the 0 is to be left alone in the denominator and further that if a zero resulting from the multiplication of a (nonzero) number by 0 is later subject to division by 0, that number is to be left unchanged. One possible explanation for these meanderings is that all the mathematicians who came after Aryabhata (though not he himself as we shall

see later) thought of 0 exclusively as the integer preceding 1 (or succeeding -1) but not also as a number that could be approached arbitrarily closely by the reciprocals of ever larger integers, or of increasing powers of 10 for that matter. The central issue here is that of limits; later on in this book we will come across other facets of this struggle with the notion of a limit, which was only resolved finally in the course of the invention of infinitesimal calculus by Madhava, when the question could no longer be evaded. This is a struggle which the Greek-European mathematical culture also went through over a very long period, beginning with Zeno and Aristotle and their paradoxes and dragging on until after the invention of calculus, before, in its turn, finding a resolution in the second half of the 18th century in the work of D'Alembert and Cauchy.

5.5 Combinatorics

Pingala does not fall in the Vedic period proper – he lived perhaps two centuries after Panini – nor combinatorics in arithmetic proper. The word *chandas* as in the title of his book *Chandaḥsūtra* has the general meaning of prosody, the science of metres. But for Panini it also meant the Vedas. The identification points to the construction of metrically regulated verse, to aid in memorisation to begin with, as another invention of the Vedic sages and bards, and is made credible by the reverence in which the Vedic metres were held. Just as the *Aṣṭādhyāyī* includes Vedic grammatical usages in its scope, so does *Chandaḥsūtra* cover the Vedic metres. So it is not inappropriate to take up a brief survey of the mathematics of the *Chandaḥsūtra* in this chapter. Its analysis of all the possible metres that can theoretically be built up from two classes of syllables, short and long, went far beyond the metres found in the Vedas and effectively froze Sanskrit prosody for all time to come, a reproach sometimes made also about Panini and grammar. In this sense, and just like the *Aṣṭādhyāyī*, *Chandaḥsūtra* marks the transition from Vedic to classical Sanskrit. There are other, more substantive parallels between the two which are best allowed to emerge as we become familiar with what Pingala accomplished.

The setting for the combinatorial problems that Pingala addressed is the fact that syllables in Sanskrit occur in two durations: short which is called *laghu* ('light') and long or *guru* ('heavy', of twice the duration of the short) in the context of prosody, symbolised by **l** and **g** in the following (Pingala himself abbreviated them to *la* and *ga* and sometimes just to *l* and *g*; they function as names of subsets of syllables, metalinguistic markers in the manner of Panini). A Sanskrit metre in its simplest manifestation is an ordered string of time durations, short or long, in one line of verse, independent of other phonetic content; abstractly it is an ordered sequence of **l**s and **g**s for a fixed (syllabic) length n. Every distinct sequence of any length (in theory) defines a distinct metre. (Several variant characterisations were considered and employed, mainly for reasons of euphony, in the long history of Indian prosody. Their classifications are specialisations or extensions of Pingala's and will not be described here).

The mathematical part of the *Chandaḥsūtra* comes as its last chapter 8, after an account of the general principles and of the metres in current use, Vedic metres in particular. The general theory that Pingala elaborates in this chapter could not possibly have been of use to poets: for a typical length of 8 syllables there are a total of 256 mathematical metres possible (the answer to question 4 below) and the vast majority of them never found any takers among even mere versifiers. One must assume that the appeal of the theory lay in the mathematical challenge (and the Indian love of classification).

The questions Pingala was concerned with are:

Q1). Provide a list, ordered in some canonical way, of all the distinct metres of length n, i.e., find a simple rule for ordering all sequences $\{s_i\}$ of the form $x_1 x_2 \cdots x_n$ in which each x_j takes one of two values, **l** or **g**.

Q2 and Q3). Given the rule, give an algorithm for the correspondence between a sequence and its position in the ordering; i.e., what is s_i for given index i, and conversely (without writing out the whole list)?

Q4). Find an efficient algorithm for the total number of sequences.

Q5). What is the number of sequences with a fixed number $k < n$ of **l** (or **g**) in them?

The *sūtra* style of Pingala being at least as terse as Panini's, we are critically dependent on his commentators in reading him as usually read, with the occasional attendant risk of misunderstanding.[14] With this proviso, here in brief are Pingala's answers to these questions.

A1). The most natural reading of the *sūtra*s giving the rule for generating the list of metres is that it is based on an elementary form of recursion. Begin with the two trivial 1-syllable metres, **g** first (I use numbers in brackets, (1), (2), etc., to index the list): (1) **g**, (2) **l**; to each, in order, adjoin on the left **g** and **l** to get the four metres of length 2: (1) **gg**, (2) **lg**, (3) **gl**, (4) **ll**. Repeat the process of adjunction of **g** and **l** now to the 2-syllable metres in the order determined by the previous step, and that results in the canonical list of 3-syllable metres: (1) **ggg**, (2) **lgg**, (3) **glg**, (4) **llg**, (5) **ggl**, (6) **lgl**, (7) **gll**, (8) **lll**. The general recursive prescription is clear: given the canonical list for length $n-1$, adjoin **g** and **l** to the left of each sequence in that list successively to generate the list for length n. This is not a particularly subtle thing to have done but it is a necessary first step since the answers to questions 2) and 3) obviously depend on the ordering. Variations in the procedure that will result in other orderings can easily be (and surely were) imagined – start with the pair (**l**, **g**) instead of (**g**, **l**), go from n to $n+1$ by adjunction on the right, etc. – but they either leave the conclusions unchanged or require only minor modifications to accommodate. In particular, if we switch **g** and **l** and adjoin on the right, then setting $\mathbf{l} = 0$ and $\mathbf{g} = 1$ – equivalently, set $\mathbf{g} = 0$ and $\mathbf{l} = 1$ and read in the reverse order as in *kaṭapayādi* – turns Pingala's ordering into

[14] For a critical approach to how to read the *Chandaḥsūtra* in the light of its many commentaries, see Jayant Shah's recent article, "A History of Piṅgala's Combinatorics", Gaṇita Bhāratī, vol. **35** (3-4) (2013).

binary place-value numbers starting with 0 ($000\cdots$) and increasing in order up to and including $2^n - 1$. This has led some to suggest that Pingala's work is the first instance in history of binary arithmetic. Actually there is no arithmetic here binary or otherwise, hardly any binary enumeration even; the analogy is no more than is to be expected in the combinatorics of 2-valued variables.

Pingala gives some prominence to the case $n = 3$: "there are 8 triples". Each triple is called a *gaṇa* ('collection') and identified, as Panini might have done, by a syllabic tag; in due course they became the subunits in terms of which a metre was often, more economically, specified (and gave rise to interesting new combinatorial problems). But it is perfectly clear that all the questions are asked and answered for general n, both from the descriptions of the algorithms and from the occasional use of words meaning "repeat" or "repeatedly". It is also clear that the building up of the list, elementary as the method is, fits the traditional understanding of a recursive algorithm as one in which the output of the $(i-1)$th application of a procedure is made the input in its ith repetition, a structural technique whose conceptual roots go down to a more remote past, to grammar, text transmission (the *padapāṭha* and the *prātiśākhya*) and indeed to the decimal number names themselves in the *Ṛgveda*.

A2) and A3). The binary count variant of Pingala's list affords a very simple answer to Q2) and Q3). The mth entry in the variant list corresponds to the binary representation of $m - 1$; so, to find the mth metre in Pingala's list, express $m - 1$ in binary form, adjoin as many 0s on its left as are required to bring its length up to n, invert the order of the resulting sequence (i.e., read it from right to left), and substitute 0 and 1 by **g** and **l** respectively. Thus, for example, the 3rd metre in Pingala's list of length 3 is obtained as $2 \to 10 \to 010 \to 010 \to$ **glg**, and the 4th metre of length 4 as $3 \to 11 \to 0011 \to 1100 \to$ **llgg**. For the converse problem of finding the index of a given metre, just reverse the arrows in the sequence of maps and add 1 to the number at the end.

There is no evidence that Pingala was familiar with the idea of binary numbers. It is in fact implicit all through his chapter 8 that the indexing was done by means of decimal numbers; among other indications, there would be no point in Q2) and Q3) otherwise. His solution, instead, is based on a recursive procedure arising from the way the list is prepared. If the index m is odd (even), the metre begins on the left with a **g** (**l**). On deleting the leftmost syllables of all the metres of Pingala's list for length n, we get a duplicated list of sequences of length $n-1$ in which the first and the second metres are the same, and so on for the following pairs. On then deleting the even-numbered sequences, we are left with Pingala's list of metres for $n-1$, to which the same procedure can be applied, as many times as necessary. In practice, the recipe goes as follows. If the index i is odd, put down **g** and update i to $(i+1)/2$; if i is even, put down **l** and update i to $i/2$. Repeat with the updated value of i, putting down the new output (**g** or **l**) to the right of the first. Continue till the string of **g**s and **l**s has length n and that gives the sequence s_i. The inverse recipe just reverses these steps.

A4). A modern student will say that the total number of metres of length n is $N_n = 2^n$ since each of the n positions can take one of two values. Pingala's way of thinking is different. His algorithm is based on taking repeated square roots, namely the factorisations $2^k = (2^{k/2})^2$ if k is even and $2^k = 2(2^{(k-1)/2})^2$ if k is odd, and applying them repeatedly starting with $k = n$ and terminating when k gets to 1, with $N_1 = 2$. The recursive mindset is evident and its manifestation more sophisticated in that i) it works in descent, from larger numbers to smaller, and is therefore guaranteed to end (as also in A2; we shall meet several examples of recursive descent, more complicated than here, in the rest of the book); and ii) there are two branches to the recursion, two different computational steps which are iterated, depending on whether the k that is (part of) the output of a given step is even or odd. It is in separating all the intermediate ks that arise along the way into even and odd subsets that Pingala resorts to the Paninian device of tagging them by *dvi* and *śūnya* respectively (see section 3 above). To take two extreme examples, the algorithm gives $N_8 = N_4^2 = N_2^4 = N_1^8$ (no surprise here) and $N_{15} = 2N_7^2 = 2(2N_3^2)^2 = 2(2(2N_1^2)^2)^2$. The *sūtra* in effect instructs us to start at the innermost level of the nested brackets and do the multiplications outwards. One can, if one wishes, think of Pingala's power algorithm as more efficient than computing 2^n directly, replacing a larger number of multiplications of 2 by itself by a smaller number of multiplications of larger numbers, but at the stage arithmetic had reached by Pingala's time, it is unlikely that it was considered an advantage: ultimately what was being computed are just moderate powers of 2. The point is sometimes made that the algorithm is the precursor of the modern computational method of finding very high powers of numbers, of use in cryptology among other applications, and that is true enough; Pingala's impetus however is certain to have been recursive rather than computational efficiency. The other somewhat strange fact is that exponentiation, as exemplified first in the powers of 10 of decimal enumeration or here in Pingala's powers of 2, was never explicitly counted among the standard arithmetical operations, then or later, except for squares and cubes.

A5). Since the variable in each position in a sequence can take only one of two values, **l** or **g**, the number $N_{n,m}$ of metres of length n which contain $m \leq n$ **l**s and $n - m$ **g**s (or *vice versa*) is the same as the number of ways in which m positions out of n can be chosen (commonly denoted in modern times by ${}_nC_m$ or ${}^nC_m$). It is also the same as the (binomial) coefficient of $x^m y^{n-m}$ (or $x^{n-m}y^m$) in the expansion of $(x+y)^n$ and its value, as we know, is

$$N_{n,m} = \frac{n(n-1)\cdots(n-m+1)}{m!} = \frac{n!}{m!(n-m)!}.$$

A long line of scholars, ancient (Indian) and modern (European and Indian) have stated with near unanimity that Pingala knew how to determine $N_{n,m}$ starting with $n = 2$ and building up, resulting in what is now called the Pascal triangle. The method was attributed to two extremely short and cryptic (even more than usual for Pingala) *sūtra*s first by a commentator of

the 10th century (Halayudha), accepted later by many. As Shah (cited above) has observed, much of the more recent controversy about exactly what Pingala proposed as the solution to Q5) has its origin in forcing a reading on these particular *sūtra*s which they cannot support. He has instead identified another *sūtra* from the same chapter which is most naturally read, with help from writers from a time much earlier than Halayudha (especially Bharata, the author of the treatise *Nāṭyaśāstra* on the performing arts, probably 2nd-1st century BCE), as giving rise to a tableau for $N_{n,m}$ which is a variant of the Pascal triangle, but which is quite distinctive in the logic of its construction; the description below is based on Shah's reading. Pingala probably started with $n = 2$ or even $n = 1$ and built up the tableau step by step, but little is added or lost if we discuss the case of general n and use descent.

The starting point is the observation that the set of all metres of length n having $m \leq n$ **l**s is obtained, without duplication, by adjoining (on the left, say) **l** to every metre of length $n - 1$ having $m - 1$ **l**s and adjoining **g** to every metre of length $n - 1$ having m **l**s (the same recursive principle as used in the indexing in A1). Therefore

$$N_{n,m} = N_{n-1,m-1} + N_{n-1,m}.$$

This is the standard 2-term recursion relation for the binomial coefficients and the Pascal triangle is just its graphical implementation beginning with $n = 2$. Later Indian authors do give this method but Pingala's original procedure, according Shah's reading of the relevant *sūtra*, was to repeat the process by applying the 2-term decomposition again to the second term on the right,

$$N_{n,m} = N_{n-1,m-1} + N_{n-2,m-1} + N_{n-2,m},$$

and iterating until the sequence terminates:

$$N_{n,m} = N_{n-1,m-1} + N_{n-2,m-1} + \cdots + N_{m-1,m-1},$$

on account of the residual term $N_{m-1,m}$ vanishing (there is no metre with more **l**s than its length). To be noted in its practical use are the special values $N_{n,n} = 1$ and $N_{n,1} = n$ which in fact follows from the recursion, as well as the symmetry $N_{n,m} = N_{n,n-m}$ (interchange **g** and **l**). It is straightforward to turn this into a modified Pascal triangle by writing the numbers $N_{n,m}$ in a plot with n and m as coordinates (see Shah).

The numbers $N_{n,m}$ are thus determined by a recursive algorithm – the strongest argument, for me, in favour of Shah's reading – which became the proptotype for many similar constructions ('sums of sums' in later terminology) in the future. They played a role in several areas of mathematics and these matters will find their due place in the rest of this book; see in particular Chapter 12.2. The combinatorial problem itself was generalised in various directions. One example was already touched upon by Pingala: defining the duration of a meter as twice the number of **g**s plus the number of **l**s, what is the number of metres of a given duration and how are **l**s and **g**s distributed in them? Later,

especially under the influence of the scientific study of music with its seven notes and four tempi of durations 1, 2, 4 and 6 units, the combinatorial issues became greatly more intricate. They were systematically tackled by another great theorist of music, Śārṅgadeva (13th century).[15] A unified mathematical treatment of several combinatorial problems, including in particular the theory of binomial coefficients, was undertaken by Narayana (Nārāyaṇa, sometimes with the honorific 'Paṇḍita' attached) in the middle of the 14th century, but still harking back to Pingala's original methods. Narayana may well have influenced Madhava who came a generation later and put the formulae for ${}_nC_m$ in the limit of large n to very productive use in the development of power series representations of trigonometric functions. These aspects, including how the recursion relation leads to the exact explicit formula for ${}_nC_m$, are discussed in Chapter 12.2 below.

In view of the sophistication of the problems Pingala posed (leave alone their answers) and their abstraction and apparent practical futility for versifiers, it has sometimes been suggested that Chapter 8 of the *Chandaḥsūtra* was a later addition. This is of course a question of great historic interest and the answer is that it is very unlikely. I have already mentioned that Bharata, who lived perhaps only a century after Pingala, echoes and elaborates the latter's combinatorial *sūtra*s, and in less opaque language; Shah, cited earlier, finds an almost perfect paraphrase in Bharata of Pingala's most futile *sūtra*, the one on the algorithm for $N_{n,m}$. It has also been pointed out[16] that the earlier chapters of the *Chandaḥsūtra* have passages which depend on the theoretical machinery of the last chapter, in fact would not make much sense without it. Altogether, there is a convincing case that the *Chandaḥsūtra* as we know it today is an organic whole and that the combinatorial theory of its last chapter can be dated around the 3rd century BCE, give or take a century.

[15]An exposition in modern mathematical terms of the ideas and methods from his *Saṅgītaratnākara* will be found in Raja Sridharan, R. Sridharan and M. D. Srinivas, "Combinatorial Methods in Indian Music: *Pratyayas* in *Saṅgītaratnākara* of Sārṅgadeva", in [SHIM].

[16]S. R. Sarma, "Śūnya in Piṅgala's *Chandaḥsūtra*" in [Sunya]. This is a useful reference for the other *pratyaya*s as well, paying special attention to the use of zero as a metalinguistic (or rather metanumeric) label for odd numbers.

Part II

The Aryabhatan Revolution

6

From 500 BCE to 500 CE

6.1 One Thousand Years of Invasions

No one disputes the fact that the writing of the *Āryabhaṭīya* in 499 CE marked a decisive turning point in the story of mathematics in India. Within its four chapters are to be found ideas and techniques of such originality and power as to have set the course, like a latter-day *Aṣṭādhyāyī*, for the enterprise of mathematics and astronomy for a long time to come. The best known of Aryabhata's mathematical achievements are the invention of plane trigonometry and, within it, the setting up and the approximate solution of the 2nd order difference equation for the functions sine and cosine, but there are also other 'gems', a term which he himself uses. From the historical perspective, as significant as these innovations was the fact that mathematics found a new driving force in astronomy.

From the early Vedic writings with their allusions to heavenly bodies, through the composition of the *Vedāṅga Jyotiṣa* and later, no connection between mathematical reasoning and astronomical phenomena was drawn, or indeed imagined, beyond the inevitable use of numbers for counting and reconciling various solar and lunar periodicities. In line with their time-keeping functions, the concerns were calendrical. With the *Āryabhaṭīya*, that concern extended to all moving heavenly bodies, the visible planets Mercury, Venus, Mars, Jupiter and Saturn in addition to the sun and the moon, and in fine detail: not just the periods, but also the details of their trajectories, changes in their positions with respect to latitude and season, their conjunctions, eclipses etc., all subjected to mathemtical analysis. A system of celestial coordinates was devised and astronomy became geometry in motion.

And, much as Aryabhata's own genius brought about this synthesis, it has a history and that history has a geopolitical background. There is very good evidence that the renewal of Indian astronomical ideas, of which the *Āryabhaṭīya* came as the high point, began some centuries earlier under the influence of Hellenestic models of planetary motion. The agency of transmission was the

P. P. Divakaran, *The Mathematics of India*, Sources and Studies in the History of Mathematics and Physical Sciences, https://doi.org/10.1007/978-981-13-1774-3_6

close and many-sided contacts between India, especially northwest India, and the world of the eastern Mediterranean in the long aftermath of the conquests of Alexander of Macedonia.

The first ever invasion of India in historical times was that of the Persian emperor Cyrus in the 6th century BCE. The occupation extended into the upper Indus basin, which had been part of the original Indo-Aryan homeland. The rest of India – that in reality meant the northern third of the country, largely settled by the Vedic people by then – does not seem to have reacted very strongly to the intrusion. But the occupied regions were exposed, for the first time since the emergence of Vedism, to fresh cultural currents, among them Persian Aramaic writing. The Kharoshti script is a direct result of this contact and writing in general may well have had its rebirth in other parts of northern India about then. Gandhara, reckoned to have been the most prosperous province of the Persian empire, gained prominence as the gateway to India. The rise of Taxila (Takṣaśila) began during this time, first probably as a trading crossroads, soon to grow into a melting pot of cultures, a cosmopolitan centre home to diverse ethnicities. A tradition of learning took root – Panini, thought to have been from Purushapura (modern Peshawar in Pakistan), very probably lived and worked under Persian rule – and flourished, especially after the Mahayana branch of Buddhism gained ascendancy. Taxila in particular grew into a great university town and we shall see that, directly or indirectly, the region of Gandhara was destined to have a major influence on the spread of new arts and sciences – astronomy and mathematics included – to other parts of northern India.

Alexander reached the Indus river in 327 BCE at the fag end of a triumphal campaign across the lands that lay between Greece and India. It is an interesting thought that his conquests included the homelands of two of the great river-basin civilisations of antiquity – Egypt and Mesopotamia – and a part of a third, the Indus Valley. There were philosophers in his entourage and, when he turned for home, he took with him Indian learned men. It is not known whether it led to any interactions between Indian and Greek scholars at that time – we are speaking about a period when the earliest *Śulbasūtra* were already five hundred years in the past and the main locus of Vedic learning had moved east to the lower Ganga basin centred around and to the north of Magadha – but within decades of Alexander's presence in India, the king of Magadha was asking one of Alexander's successors in west Asia to send him a sophist (along with some wine).

When Alexander died in Babylon on his way back from India his generals, to whom he had given charge of the conquered territories, became independent rulers. Relations among them, now kings of the far-flung domains of the erstwhile Macedonian empire, remained close, especially strong between Gandhara and Hellenistic Alexandria in Egypt: politically, economically through trade and in the sphere of the plastic arts.[1] The Indo-Greek kings, descended from

[1]For an overview of these interactions as well as for the general historical narrative, I have relied on Romila Thapar's book [Th-EI]. Several chapters in [Th-CP] cover the evidence on which these conclusions are based.

one of Alexander's generals, had their ups and downs but the line survived into the early 1st century BCE, a little after the reign of the best-known of them, Menander (Milinda in Indian records). By then the Indo-Hellenistic exchanges of ideas had become so pervasive that their effects continued to be felt in India for centuries afterwards, as can be seen in those indestructible witnesses to cultural history, namely architecture and sculpture. The Gandharan school of sculpture is still often referred to as Indo-Greek.

While this churning was going on on the western frontier, the first step in a redrawing of the political map of India was being taken at the opposite end, with its focus on Magadha. The area which is now Bihar had already, by the late Vedic period, become a centre of Vedic scholarship, the place where the final recension of the Vedas as handed down to future generations was accomplished and, along with it, the initiation of analytical linguistics (the *padapāṭha*). But the time and the place also nurtured a radical trend of vigorous opposition to orthodoxy, valuing rational thought over revealed knowledge. It was out of this ferment, but with a concern for ethical conduct added, that Buddhism and Jainism were born and it is not an accident that they were founded as institutionalised religions at the same time and in the same region. As we have seen, both played a part in pushing forward the frontier of very large numbers, but their greater and more permanent contribution to mathematical thought was to have encouraged argumentation and debate as a means of understanding the world and, hence, to have promoted logic as an autonomous discipline not subject to Vedic injunctions. Philosophical positions formalised during these intellectually turbulent times were to cast their shadow on the epistemology of mathematics, and on mathematics itself, for a long time to come.

The political event referred to above is the founding of the Maurya dynasty by Chandragupta Maurya in Magadha, within just five years of Alexander leaving India. Knowing nothing about the political arrangements of the Indus Valley people, we can say that this event marked the establishment of the first Indian empire and it changed India's cultural history for good. Chandragupta extended his domains across most of northern India and had encounters, both hostile and friendly, with the Indo-Greeks. His son and his grandson, the great and greatly admired Ashoka, built upon this foundation until, at the height of his rule, the Maurya empire covered all of India (but for the southernmost quarter) and modern Afghanistan.

After creating an empire, and perhaps to help hold it together, the Mauryan kings vastly improved the means of transport and communication, making it possible for people other than soldiers and traders and missionaries to travel long distances, and relatively quickly. Architects, sculptors, rock-carvers and other artisans went where their special competence was in demand; so did those who sought an education or had acquired one and were looking for a market for their learning. That this promoted a high level of cultural unity is evident in the uniformity of the language (Prakrit) and the script (Brahmi) of Ashoka's stone-carved edicts in all parts of the empire except the northwest

frontier. Taxila remained a centre of attraction and soon became the nursery of Mahayana Buddhist scholarship. Chandragupta Maurya, Ashoka's grandfather, was supposed to have been sent by his mentor Kauṭilya (also known as Cāṇakya) to Taxila to be trained in the skills of kingship. When the stirrings of the new astronomy began to reshape mathematical thinking in the early centuries CE, they were transmitted to distant centres fairly rapidly; Varahamihira the astronomer (and astrologer), a junior contemporary of Aryabhata and the author of *Pañcasiddhāntikā* – our prime source of information on Greek-inspired pre-Aryabhatan astronomy – was, going by the 'Mihira' (Persian for the Sun god) in his name, not of ethnic Indian ancestry but ended up in Ujjain in the north Indian plains. In a later chapter I shall put forward the possibility that Aryabhata himself, whose work was (in his own words) "lauded in Kusumapura" in Magadha, probably traced his roots to Gandhara or thereabouts.

Kautilya was also the author of a treatise on governance and statecraft, the famous *Arthaśāstra*,[2] and there are in it some incidental bits of information about the state of mathematical knowledge in the 4th century BCE. A large part of the book deals with the management of public wealth, with elaborate sections on taxation, accounts and audits, etc., which could not have been understood by the responsible officials without a fair grasp of arithmetic. Loans, mortgages and debts figure prominently and the word used for debt is *ṛṇa*, the same word that came to denote the negative (numbers) in later mathematical writing. It would seem safe to assume that the accountants were comfortable with arithmetic involving negatives.

The other noteworthy point concerns units and measures. There are very detailed specifications in the book of the progression of units of length, area, volume, weight and time that speak of a high level of standardisation, reminiscent of the Indus Valley. The standard unit of length is defined (*aṅgula*, approximately 2 cm) and it is divided eightfold, five times successively, to get an *aṇu* (particle or atom): 1 *aṅgula* = 8^5 *aṇu*. Multiples of the *aṅgula* are nowhere near as systematically organised, with what appears to be a shift from binary to decimal for larger multiples. Can it be that we are seeing here, once again, the signals of a throwback to Harappan times?

The weights progress in a sequence involving binary and decimal multiples with no special regularity. So do units of time and that is striking: there is absolutely no sign of the sixtyfold divisions of time that came to be adopted a little later (but not later than the Siddhantas and Aryabhata), first by the astronomers and then more widely. So we know for sure that the advent of sexagesimally measured time in India can be dated between 300 BCE and 300 CE say. The use of 60 in measuring time (and angles as arcs of a standard circle, but nothing else) came to India almost certainly from the Hellenistic

[2]Kauṭilya, *The Arthaśāstra*, edited and translated into English by L. N. Rangarajan, Penguin Books, New Delhi (1992). There are many other translations.

world. The lower limit on when it happened that the *Arthaśāstra* provides fits in well with the history of Indo-Greek interactions.

The Indo-Greek kings were finally uprooted in the 1st century CE as a result of the next big wave of invaders, this time from central Asia, the Shaka (Scythians). They intruded deep into western and central India and some of them and their satraps were powerful rulers with pretensions to imperial glory. The Shakas in their turn were subdued by native kings (the Satavahana), to be absorbed gradually into the Indian political and social fabric. The roughly three centuries of their prominence are important for the history of Indian numerals; much of what we know about Brahmi numbers and their evolution comes from Shaka coinage and the inscriptions of the Satavahana.

Other invasions from central Asia followed, of which the most destructive was the last in the period of our interest here. The Huna (Hūṇa; their relationship with the Huns who devastated parts of Europe and brought about the end of the Roman empire is not entirely clear), after first establishing their presence in Afghanistan, reached Gandhara and the northwestern frontier of India a little before the middle of the 5th century. But there was resistance. The north Indian plains they coveted had attained a measure of political stability and strength; out of the debris left behind by the collapse of the Maurya empire after Ashoka, there had finally emerged, in the second half of the 4th century, the imperial dynasty of the Guptas. The Guptas were also of Magadhan origin, with their main capital at Pataliputra. But the extension of the empire close to the northwestern borders and into the northern edge of the Deccan plateau saw the rise of another city, Ujjain in the Malava country, to the status almost of a second capital and it became a celebrated centre of scientific and literary activity. These times and places matter for us because it was probably during the early phase of the Gupta rule that the continuing flow of Hellenistic ideas into India became channelled into the composition of the first texts of the new astronomy, the so-called Siddhanta texts. Among several indicators of Ujjain's pivotal part in this enterprise is the prime position it came to occupy in the setting up of a coordinate system on which to base astronomical observations and computations: the zero meridian was fixed as the one passing through it. And Pataliputra itself was, or was close to, Kusumapura where Aryabhata's astronomical achievements were famously lauded.

As for the Hunas, their depradations – especially in Gandhara and especially directed against Buddhist monks and scholars – and the struggle of the Gupta kings to contain them are the backdrop against which the new astronomy and its mathematics found their definitive formulation. For that reason and because Aryabhata's own personal history may well have been entangled with the political and military events, it is worth looking at the Gupta-Huna confrontation in some detail. I will return to it after a quick look at what we know of Siddhanta astronomy and what we can learn from it of a historical nature.

6.2 The Siddhantas and the Influence of Greek Astronomy

There are only two works from which we can hope to get a picture of what the Siddhanta texts contained. The use of the past tense is deliberate. One of the two, the *Sūryasiddhānta*, is held by experts to have been heavily and often emended over a long period; though there is probably a pre-Aryabhatan core in it, most of the changes postdate him. The text that we now have is therefore not a reliable guide to the early phase of the new astronomy. But it has historical value as the work which gave European scholars their first sighting of Aryabhata's trigonometry, as parts of it came to be known and translated into English at the end of the 18th century, almost a century before *Āryabhaṭīya* itself was published. Except in that context, it will get no further mention here.

The other source is the *Pañcasiddhāntikā* of Varahamihira ([PS-K]), purportedly a compendium of earlier Siddhanta texts (" . . . I narrate . . . the secrets of astronomy in the doctrines of the teachers of the past . . . ", from the opening invocatory verse). With the reservations that go with reliance on a secondary source (which it largely is), the present section is therefore mostly concerned with what we can learn from the *Pañcasiddhāntikā* about the developments that led to Aryabhata's own work. There is a reason for the hedging and it has to do with the date of the book. We know from a later commentator on Brahmagupta that Varahamihira died in 587 CE ([PS-K], Introduction); granting him a full span of life and assuming conservatively that he was at least 25 years old when he wrote the book, its composition cannot be earlier than about 525 CE. The post-Aryabhatan dating, by a substantial 25 years or more, means that it is difficult to be sure how much his account of the early Siddhantas was coloured by knowledge of Aryabhata's own work. We shall in fact see in a moment that it is possible to guess at some of these influences and that it throws a revealing sidelight on the first use of plane trigonometry in Indian astronomy.

But, first, some basic facts about the book and its five Siddhantas. The five are *Vāsiṣṭha* (of Vasiṣṭha, the sage who already figures in the *Ṛgveda*), *Paitāmaha* (of the ancestors or, possibly, of Brahma, the god of creation who was also called *pitāmaha*), *Romaka*, *Pauliśa* (whose names, and other features detailed below, suggest strong Hellenic affiliations) and *Saura* (of the sun, more as heavenly body than divinity). They will be referred to in what follows by their initial syllables, *Vā, Pai, Ro, Pau* and *Sau* respectively. Some Siddhantas with identical names are known from later times but, according to Kuppanna Sastry, the most recent translator, editor and commentator of [PS-K] (there were others), they are not the same; the originals are lost for good. Also, though *Sau* has a name having the same meaning as the later and more theoretical *Sūryasiddhānta*, the two do not seem to have any strong causal connection ([PS-K]). The relative importance and space given to each of the five by Varahamihira is very uneven and it is to be kept in mind that what we can learn

from them is what has been filtered through by the choices he made. From the names alone, one would guess that *Pai* and *Vā* are the earliest, possibly growing out of the more primitive considerations of the *Vedāṅga Jyotiṣa* and buttressing them with numerical parameters, and that *Ro* and *Pau* represent their later transformations in the light of the new Greek inputs. That guess appears to be borne out to an extent, but there are passages in all of them which defy such a neat classification.

Most of the chapters of *Pañcasiddhāntikā* are meticulous in naming the specific Siddhantas they are based on (I will call these the named chapters) and, when two Sidhantas are covered in one chapter, in so identifying sections of that chapter. However, five of them (Chapters IV, XI, XIII - XV) (called here the independent chapters) as well as parts of Chapter I have no such affiliation. This is very useful in that we may reasonably take, as suggested by Kuppanna Sastry in [PS-K], the independent chapters as representing Varahamihira's own contribution; their contents justify that conclusion.

The word *siddhānta* can be translated somewhat freely as 'the conclusion of, or from, what has been acquired', a summing up of established knowledge. There is a theoretical bias to the sense in which the word is generally used (for example by Brahmagupta and Bhaskara II in the titles of their books) and the Siddhanta texts, making allowance for the fact that they are only partially – some of them very partially – known, fit that sense. All five are basically what later came to be known as *karaṇa* texts, manuals of procedures for the determination and prediction of the positions of the *graha* (sun, moon and the five 'star-planets' visible to the eye) as a function of time. In line with that function and as a general rule, the rationale for the procedures is not provided. Different commentators, including some modern editors, have supplied their own versions of the reasoning behind the stated results; as in many other instances (*Śulbasūtra* and *Chandaḥsūtra* among texts that we have already looked at) , such reconstructions can tell us a good deal that is historically useful.

Aside from their names, several features of *Ro* and *Pau* point to their links with Alexandria. Most directly, in both of them, the conventional origin of astronomical time (the 'epoch' $t = 0$) from which all planetary motion is evolved into the future, is taken to be mid-sunset (on a particular day) at Yavanapura. This place was (is) not in India; from the time of Alexander's invasion, Yavana or Yona (from Ionia) was the word for Greece or Greek, later loosely extended to other neighbouring locations and peoples. To compare calculations with observations made locally, with Ujjain (also called Avanti) chosen as the zero of longitude, a knowledge of the difference in local times (equivalently, longitudes) between Yavanapura and Ujjain was needed. *Pau* gives a value for this quantity. The longitude of Yavanapura determined by this value is that of Alexandria, with surprising accuracy (it is short by about 5%).[3]

[3]The verse following describes how this difference for any two places A and B is to be found, assuming known the radius R of the earth, the latitudes of A and B and the distance between them. If C is the intersection of the longitude through A and the latitude through B, and ignoring the curvature of the earth, ABC is a right triangle with AB as the hypotenuse;

Less direct evidence of a Greek influence can be seen in the method used to simplify certain fractions with large numerators and denominators to more manageable ones without serious loss of accuracy. The problem arises in the context of counting *kali* days, i.e., the number of days that have elapsed from the 'epoch', the start of a celestial cycle (*yuga*), to a given event. The fundamental issue was the precise comparison and reconciliation of time measured by two clocks with different periodicities, the moon and the sun for example. Fractions with big numerators and denominators arise in the computations needed and both *Ro* and *Pau* give good approximations which turn out to be suitably chosen truncations (the convergents) of their continued fraction expansions (see [PS-K]).

The problem of reconciliation was formulated abstractly as a linear Diophantine equation by Aryabhata and solved by him by the method known as *kuṭṭaka* (much more about it later) involving continued fraction techniques. The Siddhantas do not have such a general approach; they are only interested in simplifying particular fractions and this is achieved by a direct application of a method used by the Greeks, the Euclidean algorithm, primarily for deciding when a pair of lines is incommensurable. Because of its historical relevance and its relationship with the place-value principle (although they served very different purposes), I recall here, for positive integers rather than geometrically defined quantities, the elementary steps it involves.

Given two positive integers a and b, let $a = nb + r_1$ where n is the quotient and $r_1 < b$ is the remainder in the division a/b:

$$\frac{a}{b} = n + \frac{r_1}{b}.$$

Next, 'measure' b by r_1: $b = n_1 r_1 + r_2$ with $r_2 < r_1$,

$$\frac{a}{b} = n + \frac{1}{n_1 + \frac{r_2}{r_1}},$$

and so on, leading to the continued fraction expansion

$$\frac{a}{b} = n + \frac{1}{n_1 + \frac{1}{n_2 + \cdots}},$$

more conveniently printed as

$$\frac{a}{b} = n + \frac{1}{n_1 +}\ \frac{1}{n_2 +} \cdots .$$

AC is known from R and the latitude difference and hence BC is determined. R and the latitude of C determine the radius of the latitude circle (notice that it is R times the cosine of the latitude), which is then used to convert BC into the longitude difference. Several methods of determining the latitude are given in the texts. The question is: how was the distance between Ujjain and Alexandria found?

The fraction resulting from truncating the expansion at any stage before it terminates is a convergent; n is the first convergent, $n+1/n_1$ is the second convergent and so on. Expressed as a fraction (numerator)/(denominator), the denominators of convergents are smaller than b (the second and third convergents have n_1 and (n_1n_2+1) as the denominator for instance). What is interesting is that fractions were approximated so as to make working with them easier in this particular way when other simpler methods based on decimal place-value enumeration were to hand. It is also notable that both the Euclidean and the place-value algorithms have the same basic ingredient, the division algorithm; only the way it is iterated is different.

Considering the extent and closeness of Indo-Hellenistic contacts in the centuries separating the Siddhantas from Euclid, the injection into Indian arithmetic of a technique of Greek origin – in the Siddhantas bearing Yavana names – should not occasion much surprise. The added novelty was that Indians had no use for it in its geometric setting but, instead, turned it into a potent tool of arithmetic over the next two centuries or so, in the work of Aryabhata and Brahmagupta. While its close cousin, the place-value algorithm, was central to mathematical thought in India from the earliest Vedic period onwards, Euclid's algorithm had found no prior use; as we have seen, fractions were dealt with in a very different way in the *Śulbasūtra* approximations to $\sqrt{2}$ and π. But once the idea of dividing by the remainder, varying from step to step, became familiar, its recursive iteration would have been a very natural procedure to adopt.

There are other details pointing in the direction of Alexandria. Most prominent is the abrupt appearance of sexagesimal fractions in time keeping – a day divided into 60 units, each of them further into 60 units, and so on. As we have seen in the previous section and in Chapter 3.2, the units in which physical quantities (including time) were measured generally followed a mixture of binary and decimal sequencing for as long as we can go back and the system continued for all units except those of time until modern times. The only rational explanation for this sudden and singular change in timekeeping – which really was the main business of the Siddhantas – is that it came to India along with Alexandrian quantitative astronomy which relied on the sexagesimal system of numbers inherited from Mesopotamia. It is significant that only the measurement of time and of the arcs of a circle (angles), which is just the trajectory of a heavenly body in a day, underwent this very radical change; in particular, numbers themselves and all computations with them, even those involving fractions, remained strictly decimal. To this we may add the overlaying of a solar calendar on a previously dominant lunar one, reflected in the naming of one of the Siddhantas (*Sau*) after the sun. Just as sudden is the emergence of astrology – which is in fact the subject of Varahamihira's best-known work, the *Bṛhatsaṃhitā* – on the Indian scene.

But beyond these details, innovative as they were, the true transformation of astronomy came from the whole-hearted adoption of the Hellenic model of planetary motion based on epicycles, planets moving in circles around centres which themselves moved in circles around the earth (from which came the

notions of the mean planet and the true planet, fundamental to all astronomical computations). And along with the particular model came the more profound and universal idea of making precisely formulated mathematical models of specific, complicated, natural phenomena and the attendant power to predict events like eclipses and other conjunctions. As in Mediterranean astronomy, the models derived from the study of planetary trajectories as a function of time and thus astronomy became applied geometry or, rather, geometry became the prerequisite for doing serious astronomy. That precondition was not diffiult to meet; from Vedic times, circles and associated linear geometric objects were the defining theme of Indian geometry.

All in all, the evidence for the transformative role of the Alexandrian connection in the renewal of Indian astronomy is indisputable. It is another story that along with predictive astronomy came also, ironically, astrological forecasts – if mathematical magic can see in advance when spectacular events in the skies such as eclipses will occur, what is more natural than that the magic can be put to use to make predictions of terrestrial events and to plot the trajectory of individual lives from the personal 'epoch', the time of birth'? Yet another example of transformative new knowledge being turned into supernatural 'science' – Varahamihira himself was more celebrated as an astrologer than as an astronomer and the two streams of the study of the heavens have lived in uneasy partnership ever since.

Where and when was the material covered in the named chapters first put together? Just as obvious as the debt to Greece are the many signs of the Indian provenance of all the Siddhantas, the two Yavana ones included. Even if there has never been any doubt that they were produced in India, it is useful to summarise the evidence. First, there is the assignment of the prime longitude as that of Ujjain. More generally, with the exception of Yavanapura, Romakapura (and Romakaviṣaya, the country of Romaka) and a notional antipodes, the geography is Indian: the naming of the intersection of the 0 meridian and the equator as Lanka, the mention of Varanasi, the reference to the Himalaya mountain and the mythical Mount Meru and even India as a whole (Bhāratavarṣa). The years are counted in the *Śaka* era. The earlier teachers invoked have Indian names (Bhadraviṣṇu, Pradyumna). The technical terminology is Indian and in Sanskrit: the names of planets, zodiacal signs and constellations, even of the sexagesimal time divisions (*nāḍikā*, *vināḍikā*, *kala*). The method for determining the cardinal directions is exactly the same as that of Baudhayana's *Śulbasūtra*. The number system employed is uniformly decimal, even when dealing with sexagesimally divided units of time. And, over everything, there is the constant presence of Hindu divinities and the quite explicit association of accurate timekeeping with Hindu ritual functions. At one point, *Pau* says (in the translation of [PS-K]):

> All the injunctions of the Vedas and Smṛtis are based on the proper time, and by not performing the rites at those times the performer, especially a twice-born, acquires sin which is to be expiated. ...

That it goes on to add that studying *Ro* is a sin is an interesting sidelight; there probably were rival doctrinal schools promoted by different teachers – though they agreed on the qualitative astronomy – reminiscent of the still-to-come criticism of Aryabhata by Brahmagupta.

The scenario then is that the Siddhanta texts as known to Varahamihira emerged out of the melting pot of Hellenic and earlier Indian knowledge systems that northwest India had become during and immediately after the period of Indo-Greek rule. Indeed, because of the occurrence of dates in the *Śaka* era (with 78 CE as its 0 year) we can, with some confidence, place the genesis of the final form of the named texts to the second half of the 2nd century CE or later, allowing time for the new calendar to establish itself among astronomers and more widely. And since Varahamihira invokes the original authors anonymously as *pūrvācārya* (teachers of the past), they must have been completed well before his own time, say between mid-2nd and mid-4th century CE. The choice of Ujjain as the reference point for longitudes would suggest Malava, the country of which it was the chief city, as the most likely place of their final redaction.

If these educated conjectures about the chronology are close to the truth, we can use it to narrow down the source of the Hellenestic component of Siddhanta astronomy. Given the dates, it cannot be any other than the Alexandria of the 2nd - 4th century CE, specifically its most preeminent astronomer, Claudius Ptolemy. The chronological fit is almost perfect: Ptolemy's masterpiece, the *Almagest* (mid-2nd century), represents in final form the characteristic features of Greek astronomy that the new Indian astronomy borrowed. To the general remarks made earlier in this section, we may add, in support of this specific suggestion, the commonality of astronomical problems treated – planetary parameters relevant to the epicyclic model (mean and true motions), the emphasis on the significance of the local time, conjunctions and eclipses, etc., not to forget the emergence of horological astrology, very much in the spirit of Ptolemy's own astrology. None of this is to be found earlier in India. Above all, the idea that the motion of celestial bodies could be geometrised and that the geometry could be used to make predictions based on accurate but tedious computations from precisely made observations was new. The Siddhantas are rich in complicated calculations and they bristle with numbers, some of them very large, all expressed in a mixture of the standard number nomenclature and *bhūtasaṃkhyā*.

To add historical plausibility to the scenario, it helps to remember that, of all of Alexandria's scientific output, Ptolemy's writings were among the most widely disseminated in the ancient learned world (and hence the best preserved) and that Taxila at that time and for the two centuries or so that followed was at the height of its glory as part of that world. (There have, nevertheless, been occasional suggestions that Greek astronomical ideas travelled to India well before Ptolemy. That is highly unlikely as we will see when we come to Aryabhata's own work (Chapter 7.4)).

The mathematical skills needed to work the Ptolemaic programme through already existed in India; indeed the arithmetical insights and expertise

that could be called upon were superior to what were available in Alexandria. The sophisticated accounting methods of the *Arthaśāstra* were already 500 years in the past and the urge to explore the unboundedness of numbers exemplified by the long power lists was at its peak. In practical terms too, computation in base 10 with its reliance on memory was far easier to do than in base 60, depending as the latter did on readymade tables of squares etc. As for the geometry, what is used in the named chapters is basic and Vedic, circles and right triangles with heavy reliance on the theorem of the diagonal, descended straight from the *Śulbasūtra*, and that seems to have been enough for their purposes. There is no trigonometry as we understand the term;[4] the fruitful coming together of circle geometry and the diagonal theorem is yet to take place.

In sharp contrast, the two have come together very effectively in the independent chapters, especially in the long Chapter IV. The basic trigonometric functions sine and cosine (and versine which is 1 - cosine) are everywhere, used as though they were constructs familiar to the reader, with no definitions or explanations offered. These terms are the same as in *Āryabhaṭīya* (and in all subsequent Indian work), even to the extent of shortening the name *ardhajyā* (the half-chord = Rsine) to just *jyā*. The elementary identities satisfied by them, such as $\sin^2\theta + \cos^2\theta = 1$ (the diagonal theorem) and $\cos\theta = \sin\,(\pi/2 - \theta)$ and $\sin\theta = \cos(\pi/2 - \theta)$ (complementarity of angles), are similarly used without explanation.

More surprising is the assumed familiarity with the formula for the sine of the half-angle, $\sin^2\theta = (1 - \cos 2\theta)/2 = (1 - \sin(\pi/2 - 2\theta))/2$. Its derivation is not given in the text but we can reconstruct it from indications from later authors (Bhaskara I in his approach to the sine table, for instance). Consider an arc AB of a circle of unit radius centred at O and let C be its midpoint (Figure 6.1). If P is the intersection of OA with the chord of twice the arc AB and θ the angle defined by arc AB, then (chord) $AB = 2\sin\theta/2$, $BP = \sin\theta$ and $AP = (1 - \cos\theta)$ (AP is called *śara*, the arrow). From the diagonal theorem, $AB^2 = AP^2 + BP^2$, i.e.,

$$(2\sin\theta/2)^2 = (1 - \cos\theta)^2 + (\sin\theta)^2 = 2 - 2\cos\theta.$$

The proof is the perfect first exercise in trigonometry, an example of the simplicity and elegance that the association of suitable right triangles with arcs of a circle brought about. The key idea in it, indeed of all of trigonometry, is that what is associated with an arc is not its chord, but half the chord of twice the arc, reflected in the term *ardhajyā*. The interest of the result and its proof lies in the fact that that idea was widely and quickly grasped; once that was done, the proof itself is elementary, certainly not beyond the geometers of the time – which is one reason for my writing it out.

[4]There are two or three occurrences of the word *jyā* for sine, more as a verbal convenience than as an essential part of the argument, probably Varahamihira taking advantage of contemporary terminology while dealing with earlier ideas.

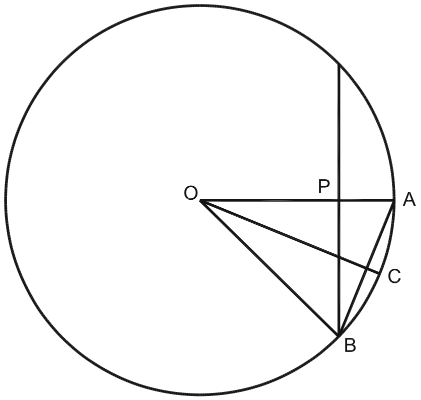

Figure 6.1: The half-angle formula

The half-angle formula is the main input in the construction of a table of sines (as distinct from one of sine differences). It is the backbone of the independent Chapter IV of which it occupies the first quarter. Sine tables have a very interesting early history in India which is best discussed in the context of the far more influential table of sine differences of Aryabhata. Here I confine myself to the table as it appears in *Pañcasiddhāntikā* and only invoke Aryabhata to the extent necessary to look at possible connections between the two. The angles for which the table was compiled are exactly those in Aryabhata's, namely 96th parts of the circle lying in the first quadrant, $\theta = (i/24)\pi/2, i = 1, 2, \cdots, 24$. The computation starts with the known values $\sin\pi/2 = 1$ and $\sin\pi/6 = 1/2$ (the latter being the statement that the side of a regular hexagon is the radius of its circumscribing circle) and applies the half-angle formula and the complementarity identity to generate the rest of the sines. As an example, the half-angle formula relates the sine of $5\pi/48$ to that of $10\pi/48$ and $20\pi/48$ which is related to the sine of $4\pi/48$ by complementarity which in turn is related to $\sin\pi/6 = 1/2$ by the half-angle formula again, thereby determining four different entries in the table from the known entry for $\pi/6$. It is an instructive exercise to convince oneself that the sine of every angle in the table can be obtained thus, and that fact is of some historical significance. Leaving the history aside for the moment, let us only note that this property of the tabular sines was, very likely, one of the original reasons for the choice of $\pi/48$ as the step size in sine tables. It then became fixed as the conventional division of the circle – note that it has no sexagesimal significance – in most future work in trigonometry. As we shall see, though Aryabhata used the same angular step in his table, he did not follow this procedure nor did he subscribe to the view that the choice was unalterable.

The half-angle formula leads directly to a table of sines rather than the sine differences of Aryabhata's table. The reason for that is well understood and not in dispute. In modern language, what he did was to set up an equation for

the second differences of the sine as a function of the angle and to 'integrate' (discrete integration, which is no more than just adding up differences) the equation once – hence the sine differences. This was a paradigm shift of major proportions with, ultimately, profound consequences for the future course of geometry and of mathematics more generally. (It is not an exaggeration to say that a good part of the rest of this book is about the working out, not excluding misunderstandings, of its consequences). In any case, it should not be unduly surprising if Varahamihira chose to avoid the conceptual novelty of the Aryabhatan approach to the computation of sines while adhering to the more conventional half-angle formula. Others, later, including Aryabhata's first influential apostle (Bhaskara I) did the same. In fact it is a possibilty that the method occurs in a work of Aryabhata himself, referred to by Brahmagupta and Bhaskara I as *Ardharātrapakṣa* (The Midnight Doctrine), now lost.

There are other minor ways in which the two tables differ, for instance in the value of the radius R of the standard circle for which the Rsines are computed as half-chords. (That may well be a remnant of Alexandrian confusions about how to measure chords in natural units, see Chapter 7.3-5). Nevertheless, given the circumstances outlined above, it becomes a persuasive hypothesis that the dramatically greater sophistication of the geometry of the independent chapters as compared to the older named Siddhantas reflects the impact of the new trigonometric ways of thinking, specifically Aryabhatan, on Indian computational astronomy. It is certain that the named Siddhantas as known to Varahamihira are pre-Aryabhtan and that the *Pañcasiddhāntikā* itself is post-Aryabhatan. Add to these facts the likelihood that the unattributed material was Varahamihira's own contribution and the case becomes compelling. Varahamihira certainly knew about Aryabhata and his work: there is a reference to him by name (in one of the independent chapters) in connection with the midnight doctrine (the vexatious issue of whether the 'epoch' is to begin at midnight or dawn at Lanka) and a criticism of him as the unnamed author ("others") of the hypothesis of the rotating earth (in another independent chapter).

As an astronomical treatise, the *Pañcasiddhāntikā* did not leave a lasting legacy; it was overtaken by the less cumbersome methods, and the more transparent model on which they are based, of the *Āryabhaṭīya*. The true historical value of the *Pañcasiddhāntikā* lies in its before-and-after role – the named chapters with no trigonometry in them and the independent chapters bristling with it – in the light it throws on the transformation of Indian astronomy, and the mathematics that went with it, brought about by Aryabhata. Once that point is taken note of, there is no room for any doubt about who invented trigonometric functions.[5]

[5]It is pertinent, as a final note, to emphasise the need for extreme caution in dating Siddhantas from different periods but known by the same name. Trigonometry occurs quite freely in the part dealing with astronomy of a Purana text (the *Viṣṇudharmottara Purāṇa*), known as *Pitāmaha Siddhānta*. As Kuppanna Sastry ([PS-K]) points out, it has little to do with the *Pai* of Varahamihira. The Purana in question cannot be earlier than the 6th century CE.

6.3 *Āryabhaṭīya* – An Overview

The *Āryabhaṭīya* is the only surviving work of Aryabhata. It is entirely in classical Sanskrit verse, consisting of 121 two-line stanzas organised into 4 parts of uneven length titled *Gītikāpāda*, *Gaṇitapāda*, *Kālakriyāpāda* and *Golapāda*. *Gītikā* (I will often omit the word *pāda*, meaning a quarter, one-fourth of the whole book) gets its name from the metre *gītikā* in which it is composed; the titles of the other chapters convey their contents: mathematics, the reckoning of time and the study of spheres (spherical astronomy and the sphere that is the earth itself, with forays into what today will be called astrophysics, very qualitative) respectively, all in the metre *āryā*. One can read the book from beginning to end in about half an hour as long as one does not attempt to absorb the full meaning; the style is terse, the sense of individual verses often obscure and, once decrypted, incisive and original. Naturally, it inspired many commentaries specifically designated as such, the last of them by Nilakantha, more than a thousand years after its writing, when original mathematical and astronomical work in India was on its last legs. Indeed it can be said truthfully that all subsequent work in astronomy and related mathematics beginning with the independent chapters of the *Pañcasiddhāntikā*, even if the books are not all titled *Āryabhaṭīyabhāṣya*, constituted a continuing and many-faceted commentary on those 121 verses. And there were only a handful of mathematical books that were not concerned with astronomy, which in itself is a measure of the impact that the *Āryabhaṭīya* had on all that came later. Aryabhata brought Greek-inspired astronomy and *Śulbasūtra* geometry together as one science, with decimal arithmetic as the binding glue, never to be separated as long as the tradition of scholarship survived.

The *Gaṇitapāda*, which is part 2, is naturally what we shall be focussing on in this book and all of the next chapter is devoted to it. The rest of the present section is meant as a quick (and inadequate) overview of Aryabhata's astronomical insights.

The first part, *Gītikā*, is a concise (10 operative stanzas) recapitulation of the essential numerical parameters characterising Aryabhata's astronomical doctrine, his own Siddhanta, as described in the three following parts; a summary of a summary as it were. It is meant to stand on its own, an *aide-mémoire* for the practising astronomer. Despite the extreme concision, several of Aryabhata's new ideas on the motion of celestial bodies (and, not to forget, earth itself) can be inferred from just these numbers. Verse 1 is an invocation (about which more in the next section) and 2 is a glossary for deciphering his peculiar system of number names that does not follow the traditional nomenclature, designed for economy of syllables – very Paninian, but it never caught on. The very first astronomical verses, 3 and 4, are already typical of the general style of the book: they list without preamble the periods (or revolutions, the term preferred by some authors, *bhagaṇa*) of the sun, the moon and the visible planets (as well as one of the theoretical constructs of the epicyclic model, the mean planets) in a *yuga*, without saying what a *yuga* is. Later, in *Kālakriyā* 7 and

8, he does say somewhat indirectly that a *yuga*, common to all the planets (*grahasāmānyaṃ yugam*), amounts to 4,320,000 solar years and the number of periods of the sun in *Gītikā* 3 matches it. The notion of a *yuga* is never otherwise, i.e., as deriving from the general properties of celestial phenomena, defined; that it is a common multiple of all the planetary periods is a conclusion we have to draw. Thus, for Aryabhata, the visible universe is periodic in time and the periodicities of the various heavenly objects, real or theoretical, assumed not to change from cycle to cycle, define his system of the world. All calculations and predictions are based on these alone, supplemented by initial conditions (the configuration at the beginning of a *yuga*) and some linear and angular dimensions (as in computing eclipses for instance).

The pinning down of the *yuga* as a unique, constant, astronomically defined interval of time – even though couched in conventional terminology like a 'day of Brahma' and so on – must have been anathema to Aryabhata's contemporaries and those who immediately followed him. In a further blow to the orthodoxy, he adds (in *Kālakriyā* 11) that time itself is without beginning and without end, *anādyanta*. The earlier system of *yuga*s, like all ritual timekeeping, was supposed to have a mythical-divine origin and the *yuga* itself was a somewhat elastic concept, tied as it was to the mythology of the creation, preservation and annihilation of the universe and the order it embodies. It is notable that, even for a brilliant and rational mathematical thinker like Brahmagupta, it is Aryabhata's messing around with time that caused the greatest unease, as evidenced by the many harsh criticisms in the *Brāhmasphuṭasiddhānta*.

The real surprise in *Gītikā* 3, however, is to see the earth included among the planets for which the numbers of periods in a *yuga* are listed. The earth was always supposed to stay still, suspended at the centre of the celestial sphere. What then is the motion which takes it through 1,582,237,500 periods in a yuga? Even before looking at the numbers, there is a strong hint that it refers to his view that the earth spins about a north-south axis from west to east; only its motion, and of no other object in the list, is qualified as eastward (*prāk*). The numbers confirm the inference: the result of dividing the number of periods of the earth in a *yuga* by that of the sun is 366.258..., which is the Aryabhatan year in (sidereal) days. *Gītikā* 6 reinforces it by giving the rate of rotaion of the earth (*bhūḥ*) as 1 minute of arc in 4 seconds. And, unlike in his revision of the concept of the *yuga*, Aryabhata is explicit in the main body of the text (*Gola* 9) that this is precisely what he meant. Here is the translation:

> From a boat moving forward, an unmoving object (*acala* which is also a synonym for 'mountain') is seen as moving backward. So are the unmoving stars (*acalāni bhāni*) seen as moving exactly west (*samapaścima*) at Lanka.

The simile – one of only four examples of a concession to literary conceits in the whole text – makes it quite clear that the verse is about more than the spin of the earth; implicit in it is his recognition that perceived motion is relative motion.

The idea of the spinning earth caused an even greater furore among astronomers than the revision of the *yuga*s which, eventually, did become part of mainstream astronomy. The earth's spin, despite sporadic support from some astronomers until about the 9th century, never did. Beginning already with Varahamihira, through Brahmagupta, and even among his own line of disciples, all kinds of objections were raised against it. The opposition led to the censoring of some of the offending material, we do not know when or by whom. *Gītikā* 6 was doctored to make it mean that it was the celestial sphere that was turning. The syllable *bhūḥ* (earth) in the original was replaced by *bham* (for the celestial sphere); though the correct text lingered on for a while in the work of a few commentators, the doctored version is what we find in all late copies and commentaries. But this (and one or two other instances of similar syllabic sleight of hand) is minor mischief when compared to what happened to the verse immediately after *Gola* 9 (the boat on the river). It was entirely rewritten in a form flatly contradicting the earth's spin: in all extant versions, *Gola* 10 goes as follows:

> The cause of rising and setting is the constant *pravaha* wind. Impelled by it, the celestial cage (or frame, *bhapañjara*, i.e., the fixed stars) and the *graha* (which includes the sun and the moon) move exactly west (*samapaścima* again) at Lanka.

That this verse 10 is a clumsy counterfeit is evident not only from its denial of verse 9 and the material in the *Gītikāpāda*, but even more from its own internal contradictions; no remotely competent astronomer could have said that both the fixed stars *and* the *graha* moved due west, never mind what the (constant) *pravaha* wind actually did. That did not prevent attempts to discredit verse 9 on the logically strange ground that it is refuted by verse 10 (Bhaṭṭotpala) or that it represents "illusory knowledge" (Parameshvara (Parameśvara) of the Nila school, a much more significant astronomer than Bhattotpala). It is a measure of the hold orthodoxy had on even the most rational savants that this fake verse is the one that survived. Nilakantha, perhaps the most philosophically aware of all Indian astronomers, is obliged to take refuge in Aryabhata's own insight: since relative motion is all that is perceived, we should leave the question at that. The original it replaced – there must have been one since the number of verses in the three substantive parts of the text remained 108, *aṣṭaśata*, which was the name by which the work was known to Brahmagupta – was lost for good.

Whether it is the earth or the celestial sphere that rotates makes no difference to Aryabhata's computational scheme, or to any earth-centred astronomical system for that matter. We must therefore put down his preoccupation with the question of what really is moving to an urge to understand natural phenomena in more than just geometric terms. There are other instances of this concern with the physical world beyond its geometric manifestation. Estimates of the diameter of the earth (in *yojana*s, a unit whose value is not precisely known) and the *graha*s are given as is the extent of the atmosphere (*kuvāyu*,

the 'air of the earth'). The *Gola* part has many speculations, some in conformity with traditional beliefs of the time, some not, but generally insightful: differences in the perception of the heavens and the earth itself by the inhabitants of the equator and the two poles, the belief that all heavenly bodies except the sun are visible because illuminated by it, and so on, as well as the extremely curious assertion that the size of the earth undergoes a very slow expansion and contraction over cosmic time. Some of these speculative ideas are also found in the *Pañcasiddhāntikā*, but in an unattributed (and hence post-Aryabhata) chapter.

The essential details of the mathematical model are in the *Kālakriyā* and the last section of the *Gola* (concerned with eclipse calculations). It is this which grew into the fine-tuned earth-centric model of the Nila school, elaborated in the *Tantrasaṃgraha* of Nilakantha and Part II of *Yuktibhāṣā*. Along the way, it underwent several revisions – the first of them by Haridatta (7th century, perhaps from Kerala) and the last by Parameshvara (15th century) – forced by better observational data and better values for the basic parameters. But the theoretical foundations remained unchanged until the end. The *Kālakriyā* also has surprising philosophical asides, such as the remark on the nature of time already cited. After saying that time itself is without beginning or end, Aryabhata adds that it is by means of the heaveny bodies that we are enabled to measure it: the celestial sphere is a universal clock whose dial is marked by the fixed stars and whose moving hands are the *graha*.

6.4 Who was Aryabhata?

There are very few things we know about Aryabhata the person. We know definitely that his name was Āryabhaṭa, not Āryabhaṭṭa as some have tended to spell it. We know it, first, because that is how he spelt it himself, in three places in the book – two invocatory stanzas, *Gītikā* 1 and *Gaṇita* 1, and the very last verse in which he refers to his book as having the name *Āryabhaṭīya* (while cursing those who would imitate or falsify him). That is also how all the astronomers who came after referred to him. Besides, changing *ṭa* to *ṭṭa* will make the preceding *bha* a long syllable and verses in which the name occurs will no longer scan. This is not as trivial an issue as it may seem. Bhaṭṭa would make him a Brahmin who is a teacher by profession. Bhaṭa has various meanings, a common one of the time being a guardian of a fort or a soldier. That would not definitely rule out the possibility that Aryabhata was a Brahmin; not every learned Brahmin had an identifiable Brahmin name. There was one other learned *bhaṭa* we know of, the celebrated physician Vāgbhaṭa, whose life and times very likely overlapped with Aryabhata's and who is credited with the founding of the school of Ayurvedic medicine that later thrived in Kerala. It is generally thought, certainly in Kerala, that Vagbhata was not a Brahmin but a Buddhist.

From the text we know two more things about him: that he lived in a place named Kusumapura – "the knowledge that was lauded in Kusumapura" (*Gaṇita* 1) – and that he was 23 years old when the work was composed in 499 CE (in the commonly accepted reading). This last occurs as an example of time-reckoning (*Kālakriyā* 10) and actually says no more than that his 23rd year coincided with the 3,600th year of the then (and still) current (quarter) *yuga* from whose advent in 3,102 BCE the *kali* days are counted. Strictly speaking therefore, it only fixes the year of his birth, not the year of the writing of the *Āryabhaṭīya*. Nevertheless, almost all commentators, ancient and modern, date the book to 499. As for Kusumapura, the general consensus is that it was, or was close to, Pataliputra (modern Patna), on the basis of Bhaskara I's identification of the place and the only historical reference to it, the legend of the founding of the Magadhan capital in the Chinese pilgrim-scholar Huen-Tsang's record of his travels in India (7th century). Though it is described as a place full of sangharamas (a word Huen-Tsang employs for monastic retreats and centres of learning), it was almost certainly not Nalanda, the preeminent Mahayana 'university town' of Magadha; the descriptions of the two places are separated by a long and complicated itinerary in the Chinese scholar's journal. Magadha and eastern-central India more generally saw the founding of many academic centres, mostly but not exclusively Mahayana Buddhist, at the turn of the 5th-6th centuries and later. Huen-Tsang counts hundreds of sangharamas and tens of thousands of monks in the major centres he visited and there is an excellent historical reason for their sudden proliferation which we shall come to in a moment. In any case, in the absence of more definite information, it is not unreasonable to accept that Aryabhata lived and worked in Magadha and his boast about his work being honoured there would seem to suggest that his Kusumapura was an important centre of study and research.[6]

Of his numerous followers and commentators, only two have anything pertinent to add to this meagre stock of personal information. Bhaskara I, about a century after Aryabhata, refers to him as *aśmakīya*, of or from the country of Ashmaka, and Nilakantha (a thousand years later) says in his *Āryabhaṭīyabhāṣya* that he was *aśmakajanapadajāta*, born in the republic of Ashmaka. Nilakantha sometimes uses *janapada* to mean simply a country or a state and perhaps he was just citing Bhaskara I who, it may be added, was influential in Kerala already before Madhava and the founding of the Nila line; in any case, the once flourishing republican states of north India had more or less disappeared by the time Aryabhata was born. At the least, we know that he was not a native son of Magadha. And that, aside from one or two vague allusions to his having occupied some position of seniority or honour, is the sum total of the information we have from textual sources about him as a living person rather than as a disembodied genius; the mystery of the man remains.

[6]But even in this, we have to be cautious. Huen-Tsang visited another Kusumapura in his travels, far from Magadha in the country of Kanyakubja (Kannauj), also a place of learning with many monasteries. But he was writing about a time more than a century after Aryabhata, when Kannauj was the capital of Harsha's kingdom.

About Ashmaka we are on slightly more solid ground. Aśmaka (or Assaka, its Prakrit orthography, or Aśvaka) was one of the early republics and it figures in many texts beginning with Panini and continuing on to the Buddhist canonical literature and the epics. There were however two Ashmakas, historically and geographically distinct, one neighbouring Gandhara and the other in central India, the northwestern part of the peninsula between the rivers Narmada and Godavari. One Buddhist text (*Digganikāya*) in fact mentions a migration during the lifetime of the Buddha from the northwestern to the central Indian Ashmaka. What is certain, and very relevant to our concerns, is that in the 5th century CE, at a time when the Gupta empire was still powerful, there was an exodus of Mahayana Buddhists from Gandhara and contiguous areas to the same trans-Narmada region. A branch of this second migration travelled also to the east, to the Buddha's own homeland and, in due course, farther afield to Bengal and Orissa. It is against this backdrop, of the displacement of a people who counted among them some of the most accomplished scholars of the time – one of whom, Asanga, mentions an Assaka country in the Indus basin – and the political and military events that caused it, that Aryabhata's life and work played out. An understanding of these events and their interconnectedness not only serves as a case study of how the forces of history promoted the creation and sustenance, over a vast geographical area, of a culture which we recognise as quintessentially Indian but also, I will argue, throw some light on the otherwise obscure story of Aryabhata himself. As we have already seen, the 5th century was not the first time that the flow of cultural currents over long distances shaped Indian mathematics nor was it destined to be the last.

First, a quick glance at the political happenings. In the first half of the 5th century, the sway of the Gupta empire extended over almost the whole of northern India and part of modern Afghanistan. Towards the middle of the century the Hunas, who had earlier installed themselves in northern Afghanistan, moved south and overran and devastated most of northwestern India, Gandhara with particular ferocity. The Chinese traveller Sung-Yun who passed through the region in the year 520 has left a graphic description of what he saw and heard about the "barbarous atrocities" wrought by the Hunas two generations earlier. The most vicious attacks were targeted at Buddhism and Buddhist scholarship (the Huna ruler of Gandhara "did not believe the law of the Buddha", Sung-Yun). Monasteries, study centres and libraries were razed. About 450 CE, Taxila as a vibrant crossroads of learning and as the most important centre of Mahayana Buddhism ceased to exist, never to rise again, its ruins an eloquent confirmation of its past glory (and of the veracity of Sung-Yun's account). What happened to the population? Where did the uprooted monks and the professors and the students go?

Despite the extent and power of the Gupta empire, written records that might have helped reconstruct the political and human history of the period are relatively sparse. Even the succession of kings, their names and their regnal dates are subjects on which there are different views among historians. So, once again, we are forced to fall back on the permanent legacy bequeathed by them

in the form of architecture and sculpture. Here we are fortunate. There is an enormous amount of Gupta art – of an extraordinary beauty, it may be added – and it is found in virtually all parts of the land where they had influence, from Mathura in the north, to Sarnath in Magadha to the east, to Ajanta and Kanheri (close to Mumbai) to the south and west, to mention only the richest and most important sites.

To look at Gupta sculpture of the latter half of the 5th century with an eye attuned to late Gandharan Mahayana art is to see a thematic and plastic unity that is quite startling.[7] The dates are already suggestive: the first Mahayana cave temples of Ajanta – which was an important Hinayana site until the 1st century CE – were excavated between 460 and 465. Karla and Nasik (about 150 and 200 kilometers from Mumbai) as well as Kanheri, which were also Hinayana sites around the beginning of the common era, saw a revival of Buddhism, but now of the Mahayana variety. The sudden flowering of Mahayana Buddhist architecture and sculpture in faraway Sarnath and other locations in eastern India also dates from the same time. The defining characteristics of this new Buddhist art can all be traced back to late Gandhara in theme, style and symbolism. A particularly noteworthy feature – among many such (see Divakaran, cited above) – is the great popularity of the representations of the Buddha as Amitābha, the 'lord of infinite radiance'. The transfiguration of the historical Buddha into the divine and transcendental Amitabha took place in Gandhara and his symbolic affiliation with the sun was to an extent the result of Persian influence. In the philosophy of the dominant Mahayana sect of late Gandhara, he presided over an infinity of universes replicated in space and time; most of the rewriting of old Buddhist texts to reflect the newly powerful Mahayana faith, *Lalitavistāra* with its unending lists of powers of 10 for example (see Chapter 5.3), was done in Gandhara.

The natural explanation for the renovation and reoccupation of old Buddhist sites as Mahayana sanctuaries with a strong Gandharan signature (and for the establishment of new ones), all at the safe peripheries of the empire, is that these are the places at which the community displaced from Gandhara found refuge. What little we know from recorded history supports such a view. The scale of the resettlement programme and the magnificence of the newly refurbished monastic complexes could not have been possible without high patronage, as attested by several inscriptions in Ajanta (which fell in the territory of powerful feudatories of the Guptas) and elsewhere. The royal personages responsible for these massive operations was most likely king Skandagupta who had defeated the Hunas somewhere in the north Indian plains around 455 (the Hunas rose again later but that is another story) or his feudatories and successors. The royal family, staunch Hindus earlier, suddenly began to adopt

[7] The art-historical material summarised here is from Odile Divakaran, "Avalokiteśvara – from the North-West to the Western Caves", East and West, vol. **39** (1989), p. 145, where it is developed in much greater detail. An analytic account of the interplay of politics, faith and art during those troubled times will also be found there.

Buddhist names or titles: Skandagupta became Shakraditya and two of his successors were named Buddhagupta and Tathagatagupta.

More to the point for us, the consequences of the devastation of Gandhara were not confined just to the spread of the faith to the margins of the shrinking Gupta empire; it also shifted the intellectual centre of gravity eastwards. As though to compensate for the loss of Taxila, study centres, some of them major universities, sprang up in eastern and central India, the most famous being of course Nalanda. According to Huen-Tsang, it was founded by king Shakraditya and subsequently enlarged and embellished by his successors beginning with his son Buddhagupta. That identifies Shakraditya as Skandagupta who had a short rule, ca. 455 - 470, and hence fixes the founding of Nalanda as within 15 years of the destruction of Taxila. The monastery at Nalanda was headed for some time by Vasubandhu, the great Mahayana spiritual master and philosopher (and brother of another, Asanga) originally from Purushapura (Peshawar) in Gandhara.

The second half of the 5th century was thus a period of great political turmoil and, at the same time, one of intellectual and cultural renewal. The impetus for the renewal came largely from the north-west; both its historical background and its Mahayana colouring argue for that. And it is in the middle of this disruptive half-century that Aryabhata the *aśmakīya* was born we do not know where, and lived and did his work in Magadha, a younger contemporary of king Buddhagupta, benefactor of Nalanda and other centres of learning. What is more natural then than that Aryabhata's Ashmaka, the original domicile of his family – Bhaskara I's *aśmakīya* need not mean anything more than that – was in fact in Gandhara or thereabouts?

Was Aryabhata a Buddhist himself? There is nothing in the *Āryabhaṭīya* that compels that conclusion. There is also nothing, nothing at all, to suggest that he was an orthodox Hindu of his times or, for that matter, a Jaina. That may come as something of a surprise and so it is useful to take a closer look at what the text itself offers by way of evidence (and not rely on the words of later commentators; we have seen how misleading that can sometimes be). In technical texts, one place to look for the authors' preferences in the matter of faith is in the invocatory verses, offering devotion to particular gods and seeking their blessings. Both Varahamihira and Bhaskara I invoke the grace of their favourite deity Sūrya – Mihira and Bhaskara are both names of the Sun god. This was a universal convention observed until the very end; Bhaskara II, typically, paid obeisance to a generous collection of divinities and Nilakantha begins and ends *Tantrasaṃgraha* by eulogising Viṣṇu. Exceptions to the rule are rare enough to stand out: before launching into Malayalam prose, *Yuktibhāṣā* has a short verse in Sanskrit in which Jyeshthadeva bows at the "lotus feet" of his guru (probably Nilakantha) but, interestingly, there is no mention of any god.

The *Āryabhaṭīya* has three benedictory verses, the first stanzas of the *Gītikā* and the *Gaṇita* and the last-but-one stanza of the whole book, *Gola* 49, and one (*Gītikā* 13) about the benefits to be derived from a knowledge of

the *Gītikā*. All traditional commentators, for understandable reasons, and most modern ones have taken the name *brahma* that occurs in all of them as meaning the relatively late Hindu god of creation Brahmā (with a long *a* at the end). But *brahma* (stem *brahman*) has another and older meaning going back to the later Books of the *Ṛgveda*, not as a particularised god, but as an abstraction, a pervasive spiritual and philosophical presence that is both individual and universal. Which of these did Aryabhata have in mind?

Gītikā 1 offers *praṇāma* to *kaṃ satyāṃ devatāṃ paraṃ brahma* who is *ekam* and *anekam*. The word *devata* as distinct from *deva* (god) has the more specific meaning 'divinity' or 'godhead' and an accurate rendering of *param* will be as 'transcendental'; so a faithful translation would be: "to Ka the supreme *brahma* who is one and not one, the true divinity". *Gītikā* 13 says that knowledge of the *Gītikāpāda* leads to the attainment of *paraṃ brahma* after penetrating the orbits of the *graha* and the celestial sphere. *Gaṇita* 1 (the Kusumapura stanza) bows to *brahma* followed by all the *graha*, mentioned individually by name, then to the collection of fixed stars. Finally we have *Gola* 49 which I quote in full:

> From the ocean of true and false knowledge, by the grace of *brahma*, the best of gems that is true knowledge has been brought up by me, by immersing [myself in it] and by [means of] the boat of my own intelligence (*svamatināvā*).

These verses have several layers of significance. The least ambiguous concerns the gender of the noun *brahma*, something that can be decided directly from the case endings. The first three occurrences are in the accusative case and the endings in the text are right only for the neuter gender, while the last (genitive) is consistent with both masculine and neuter. The anthropomorphic (even allowing for his four – or five – heads) Brahmā is, equally certainly, a masculine god. There is no doubt about who – or which – Aryabhata's *brahma* was.

At a more subtle level, *brahma* is invoked in two of these verses along with the planets and the celestial sphere, but as a concept above and beyond the visible material objects of the universe or, perhaps, as the very essence of that universe. This is a far cry from the notion of Brahmā prevalent in the Hinduism of Aryabhata's time. It harks back to the agnostic and speculative spirit of the later Upanishads and, eventually, to some of the late poems of the *Ṛgveda*, especially of its Book X. The pairing of Ka and *brahma* in the very first verse – it is in fact reasonable to read *Gītikā* 1 as implying the identity of the two – strengthens this interpretation. Ka first occurs in late *Ṛgveda* and literally means 'Who': in some of the philosophical poems of Book X, for example the oft-quoted poems 121 and 129 about creation, existence and causation, it is both the question (who is?) and its answer (Who is), agnosticism taken to its logical limit. There is no other Ka in the Hindu pantheon though, as often in India, later literature sometimes identified Ka with Prajāpati the first progenitor who ultimately morphed into Brahmā.

The picture we get of Aryabhata's personality is then one of fierce independence of mind, in his scientific thinking as much as in a radical world-view that he was not hesitant to profess. The verse *Gola* 49 quoted earlier makes it clear: not only does he disown any indebtedness to supernatural revelation or inspiration – apart from the acknowledgement of the grace of *brahma* as the sustaining spirit of the cosmos – it is to his own intellect, *sva-mati*, that he credits the "best of gems" of astronomical knowledge.[8] To come back to the question we started with: did he have a personal faith and, if yes, what was it? Aside from the circumstantial evidence – his probable ancestral origins in Gandhara, the spread of Mahayana piety and the intellectual life it nourished in the great centres of learning under the later Gupta kings – there is almost nothing in the *Āryabhaṭīya* or in writings on it to suggest that he was a Buddhist; 'almost' because there are two incidental allusions in the text, not essential to the astronomy, which may point to a Mahayana worldview. In *Gola* 11, the Meru mountain, the *axis mundi* through the centre of the earth, is said to be surrounded by a radiant garden made of jewels and in *Gola* 12, it is identified as the paradise inhabited by the immortals. It corresponds almost perfectly with the bejewelled Mahayana paradise of the immortals called *Sukhāvatī*, presided over by Amitabha, of which many representations exist in late Gandharan sculpture. The other indication is the insistence in *Kālakriyā* 11 that, though planetary orbits are periodic in time, time itself has neither beginning nor end; as we have had occasion to note, an infinite but cyclic universe of which the transcendental Amitabha was the presiding divinity was a cornerstone of late Mahayana cosmogony. The idea would have been anathema to orthodox Hindus of the time; the open-mindedness and scepticism of the late Vedic hymns had by then been overtaken by a mythical view of the origin of things in a supernatural act of creation or a sequence of such acts. The virulence of Brahmagupta's attacks and the complicity of Aryabhata's own intellectual lineage in the rewriting of offensive passages are proof of the antagonism he must have provoked.

In the end, the question does not matter a great deal. By the time of Aryabhata's life, Mahayana Buddhism was itself at the start of a momentous transformation, one of rapid Hinduisation and consequent loss of identity that, in course of time, contributed to its disappearance from most of India. In the centuries that followed, there was only one prominent mathematician, in the south of India, who was not a Hindu (going by the invocatory passages and, of course, his name), the Jaina Mahāvīra in the 9th century. It has also been suggested that many men of learning, astronomers among them, really were adherents of philosophical systems such as *lokāyata* or *sāṃkhya* which were frowned upon by mainstream Hinduism. Perhaps Jyeshthadeva, whose *Yuktibhāṣā* bows to no god, was one such.

---

[8] The next verse, the last of the book, somewhat spoils the effect; it speaks of *Āryabhaṭīya* being the same as the science which, depending on the reading, had earlier come into existence by itself or was due to one who was self-created. Given the history of mutilation the text was subjected to, we have to wonder how authentic it is.

But all this was still to come. Historians often refer to the Gupta period as a golden age for culture and the arts. For the sciences too the 5th-6th centuries were a time of renewal; Aryabhata was part of that renaissance as were, of those we have had occasion to mention, Vagbhata and Bhartrhari.

6.5 The Bakhshali Manuscript: Where Does it Fit in?

In deciding to take up the Bakhshali manuscript as the concluding section of the present chapter, I have taken a position regarding its dating. The reasons are best allowed to emerge as we take a look at its contents and style and compare them to the dated *Āryabhaṭīya*.

The facts about the physical book are well known. It is written in ink, probably with a brush made from a fibrous twig, on the dried bark of the Himalayan birch tree. What exists now is a collection of 70 'folios' written on both sides, preserved between sheets of mica in the Bodleian Library of Oxford University. It is definitely incomplete and it is not known how many folios are missing. There is however a colophon giving some personal details about the author (though not his name nor a name for the book). Such data are normally found towards the end of a book and it is therefore reasonable to suppose that the missing part did not have material that logically followed from what is in the surviving sheets; it is unlikely that anything significantly more advanced than what we can learn from them has been lost. The fact remains, nevertheless, that anything said about the book (for example, that it has little geometry) is tentative to the extent that it is based on partial information.

The history of the manuscript is also well known. It was dug up accidentally in village Bakhshali, some 150 km northwest of Taxila and 80 km northeast of Peshawar, where it seems to have been buried in some distant past. The manuscript ended up in Calcutta with the indologist A. F. R. Hoernle who lectured and wrote about it. One other person who had access to it and published its first translation in English with a commentary was G. R. Kaye, who had also written about other topics in early Indian mathematics, in particular the *Śulbasūtra*. Kaye's views were coloured by such a strong colonial bias that, after a period of vigorous refutation, most scholars paid little attention to them, a precedent I shall follow. Fortunately, we now have the definitive study of the manuscript in Takao Hayashi's critical edition in English with comprehensive annotations ([BM-H]); much of what follows in this section will be best appreciated with Hayashi to hand.

The Bakhshali manuscript is primarily a work on arithmetic and practical algebra (by which is meant problems in arithmetic which are most conveniently solved by introducing a symbol for the unknown answer and treating it like a number). It is the very first treatise devoted to arithmetical questions, independently of its date within a reasonable range, even if it is as late as Hayashi's

preferred 7th century; Śrīdhara's *Pāṭīgaṇita* is not earlier than the 8th and Mahavira is another century later. It should not be the least bit surprising then if we find in it concepts and methods which are not to be found in earlier texts – to belabour the point, there was no earlier text to trace them back to – except, if we are lucky, in passing, as the need arose. In other words, the novelty of its content and treatment is not an inverse measure of its antiquity and cannot logically be used to rule out an early date. As emphasised in the Introduction, the true historical significance of the book lies in whether we can assign a reasonably precise date for it and that cannot be done by looking for parallels with Brahmagupta or Bhaskara I. Sadly, there is no positive evidence which will determine it; all we can do is to look for suggestive indications.

First and foremost, the manuscript employs a full-fledged symbolic decimal notation for numbers (with the highest place to the left). Several of the symbols for the atomic numerals are close cousins of the Brahmi numbers of the rock inscriptions of the 2nd-4th centuries. Some (for example, 7) are identical. Others are easily explained as arising from the ligatures which are inevitably introduced when going from incisions on stone using a chisel to handwriting with a brush; for instance the Bakhshali 3, which is the same as in Devanagari which in turn is almost the same as the 3 printed here, comes out of the three horizontal bars of the rock-cut Brahmi $\equiv$ by starting at the left end of the top bar and ending at the left end of the bottom bar, all in one brush stroke, as we write today. The overall impression conveyed is that, while borrowing and adapting the Brahmi numbers signs, Bakhshali had overcome the confusions Brahmi numbers suffered from in trying to symbolically represent the prevalent standard number names (see Chapter 5.3). That was a significant step and would support the idea that Bakhshali was composed after the last Brahmi inscriptions were carved, say after 300 CE. But how much later? In some of the narrative passages, it adopts the practice of naming numbers by naming groups of numerals in sequence, as we do today by reading 1947 for instance as nineteen-fortyseven; an example of the coexistence of the naming paradigm with the more recent positional writing? There are only two or three occurrences of *bhūtasaṃkhyā*, in sharp contrast to the Siddhanta texts.

And there is a symbol for zero. Zeros occur very freely, in the middle as well as the extreme right positions in numbers, and the symbol is invariably a dot, the *bindu*. That is another 'first', the first written zero in India, even if the manuscript is considered to be late; pushing it to the 7th century will make the Bakhshali *bindu* roughly contemporaneous with the southeast Asian written zeros, which come both as a dot and a circle. Does this mean that the evolution of the dot into the circle is still in the future at the time of Bakhshali? We have already seen (Chapter 5.3) that the word *śūnyabindu*, a dot denoting zero, occurs in literature as early as the 4th century CE.

The dot occurs also in another guise. When it is by itself, not as a numeral entry in a number having more than one place, it stands for the (temporarily unknown) solution in a problem that is algebraically formulated, the x in, say, a quadratic equation. The interesting point here is the symbolic pairing of the

unknown x with zero and it points to the way place-entries in decimal numbers were thought about. The analogy is that, to begin with, the numeral that will occupy a particular place is supposed to be undecided and the place itself supposed vacant or empty, zero by default. The assigning of a numeral to a place results in the occupation of that place by that numeral; a late text like *Yuktibhāṣā* still uses a word which would normally mean 'invasion' for this operation of occupation. The underlying point of view is global and abstract: start with a sequence of vacant places (the powers of 10) and choose any one of the numerals 0 to 9 to occupy each of them (the coefficients in the polynomial representation) to generate all numbers; the dot for the unknown similarly denotes a vacancy which is to be occupied by a definite number once the problem is solved. For decimal numbers, the idea was first described by Bhaskara I but, again, it may very well go back to their first written representation and so is not of much use in dating Bakhshali.

The basic arithmetical operations are dealt with with the same facility as the numbers are, as though familiar from long usage. The originality is in the way they are expressed in writing, by means of symbols which are the first syllables of the words describing the corresponding operations (or, sometimes, variations on them). Thus, generally, **yu** (from *yuta*) is addition, **gu** (from *guṇita*) is multiplication, **mū** (from *varga-mūla*) is the square root (which is its literal meaning): a **yu** $b = a + b$, ab **gu** $= a \times b$, a **mū** $= \sqrt{a}$, etc. In effect they function as abstract symbols. Similarly, though serving a different purpose, *yā* (from *yāvat-tāvat*, 'as much as is') is the name for the unknown in an algebraic problem, which it remained in the more formal algebraic considerations of Brahmagupta and Bhaskara II. It hardly needs to be pointed out that there was a model for these conventional devices, namely Panini's metalinguistic labels. The identification of the symbols with mathematical operations is aided by the fact that problems are posed, sums worked out and equations written using a calculational schema and, at the same time, they are also stated in narrative form: it is like working with a bilingual text or a glossary like Panini's *paribhāṣā* section.

The schema is unique to Bakhshali, though there are indications, in Sridhara for example, that similar schematic methods were used in practical computation 'behind the scenes' or in face-to-face teaching.

The textual material comes in two styles, *sūtra*s in which are given general rules and constructions, and commentaries applying the rules to concrete problems. (As is often the case, the division of the material into the two categories is not watertight). Some *sūtra*s are called quoted *sūtra*s by Hayashi; he considers them to have had an independent existence because they are cited more than once in the text. In fact several modern commentators tend to the view that the *sūtra*s and the commentary may not be by the same author nor from the same period. If that is true, it will make the positioning of the manuscript in the line of development of early Indian mathematics even more problematic than it otherwise is. On the other hand, authors providing their own explanatory notes to a more tersely written original, sometimes after the passage of years, are not

unknown: Bhartrhari (the *vṛtti* to *Vākyapadīya*) and Bhaskara II (the *Vāsanā* to *Siddhāntaśiromaṇi*) possibly, Nilakantha (*Siddhāntadarpaṇa*) definitely.

Moving on to the more substantive mathematical aspects, the arithmetical and the algebraic culture of the Bakhshali manuscript is roughly of the same level of accomplishment as the corresponding material in the *Āryabhaṭīya*, to pick the earliest such parallel, but (of course) less tersely described. Fractions and negative numbers occur routinely. Curiously, whole numbers are sometimes written as fractions with 1 as the denominator. Should one see in this a recognition, not fully crystallised, that integers are special kinds of rationals? Some of the general principles are explicitly described, some not, a lapse which many others including Aryabhata are guilty of. Among the rules written down in the *sūtra*s (details of concrete problems posed and solved can be found in [BM-H]) are $a \pm (-b) = a \mp b$, $(a/b) \times (c/d) = (ac/bd)$ and the corresponding rule for division, but the rule for adding fractions by finding the common denominator is not given – maybe it was in the lost folios. Several varieties of linear equations are treated, including systems with more unknowns than equations (indeterminate or underdetermined systems). The solution of the general quadratic equation in one unknown is known and used and underdetermined quadratics of the form $xy = ax + by$ make an appearance. Finally, the formula for the sum of an arithmetic series of n terms with first term a and step d is given:

$$S = \left(\frac{(n-1)d}{2} + a \right) n.$$

Some of these are found explicitly stated in the *Āryabhaṭīya*, and a knowledge of the others is implicit; more to the point, several of them are developed more fully by Aryabhata. Thus, the rule for multiplying and dividing fractions is followed in the same stanza by the one for their addition and subtraction by reduction to a common denominator (*Gaṇita* 27). The next two verses of *Gaṇita* then lay down the general method of solving linear equations involving the four operations by inverting the order of the operations, transposition etc., for one variable and for several. The mathematics is of course elementary; what is striking is that Aryabhata lays out the principles in complete generality, not through numerical examples. Similarly, the formula for the sum of the arithmetic series is supplemented in the *Āryabhaṭīya* by the converse problem of finding the number of terms given the other data, which requires the solution of a quadratic equation. The special case $a = 1, d = 1$: $S_1(n) = \Sigma_{i=1}^{n} i$ (the sum of the natural numbers) is taken to have been understood because the next few verses give the "sum of sums" $S_2(n) = \Sigma_{i=1}^{n} S_1(i)$ as well as the sums of squares and cubes.

The point about concentrating on the *Āryabhaṭīya* as a benchmark to compare the mathematics of Bakhshali with is the following. In order to determine a plausible upper limit date for a text A, it is not enough just to look at another (dated) text B with which it has some material in common. The comparison will be much more decisive, though not necessarily conclusive, if such common elements are significantly better developed in B than in A and,

conversely, if A has nothing that is more advanced than in B; the earliest such B can then be considered seriously as having been written after A. Does the *Āryabhaṭīya* meet this criterion in relation to Bakhshali? As a working hypothesis, the answer would seem to be positive. To summarise the evidence, Aryabhata's account of linear equations is more abstract and general and more comprehensive. Bakhshali has the basic arithmetic series, but it stops there; Aryabhata continues on to the more demanding task of summing squares and cubes of natural numbers. Most tellingly, though Bakhshali is not unfamiliar with Diophantine problems, it has no coherent point of view (as far as we can tell from the surviving parts), nothing comparable with the generality of Aryabhata's *kuṭṭaka* method of solving the linear Diophantine equation in two unknowns.

A problem treated by both the Bakhshali manuscript and the *Āryabhaṭīya*, but from intrinsically different viewpoints, is that of finding the square root of a (positive) number n. Aryabhata's prescription is incomplete and the relevant verse (*Gaṇita* 4) is ambiguous to interpret and implement. More pertinent for us is that it is an algorithm tailored to the decimal representation of n and hence difficult to capture abstractly in a compact formula valid for all n – commentators are usually content to illustrate it for particular numbers. The Bakhshali recipe is base-independent and fully algebraic, and it is made clear that it is approximate and meant to be used for a (any) non-square number n. It occurs as one of the quoted *sūtra*s (Hayashi's Q2) and transcribes into modern notation as the formula

$$\sqrt{n} = m + \frac{r}{2m} - \frac{1}{2}\left(\frac{r}{2m}\right)^2 \frac{1}{m + \frac{r}{2m}},$$

where $m^2 < n$ is the square number nearest to n and $r = n - m^2$. Before turning to the question of what the greater generality of the prescription may mean in terms of priority, let us take a look at how it may have been arrived at. That will also help throw light on one or two other points concerning its history.

Once a choice of the first approximation m is made (common sense tells us that the optimal choice is to minimise $n - m^2$, but the wording of the text is not unambiguous), the formula becomes an estimate of the unknown error $\sqrt{n} - m$ as a rational function of the known $r = n - m^2$. Together, these two steps fit into the general philosophy of iterated refining that pervades Indian mathematics of all periods (and known to the much later *Yuktibhāṣā* under the rubric *saṃskāram*). We have already seen how it works in the extraction of the *Śulbasūtra* square root of 2 (Chapter 2.5). Its extension to general n is straightforward. Here are the details.

The first guess is obviously $(\sqrt{n})_1 = m$. The equation $n = m^2 + r$ can be rewritten

$$n = \left(m + \frac{r}{2m}\right)^2 - \left(\frac{r}{2m}\right)^2,$$

a type of reordering familiar from the early days of the geometric algebra of squares and rectangles of the *Śulbasūtra*. Neglecting the second term, the second

approximation is

$$(\sqrt{n})_2 = m + \frac{r}{2m}.$$

Next, define r' to be the error in $(\sqrt{n})_2$:

$$(m^2 + r)^{\frac{1}{2}} = m + \frac{r}{2m} + r'.$$

Squaring and neglecting r'^2 leads to

$$r' = -\left(\frac{r}{2m}\right)^2 \frac{1}{2\left(m + \frac{r}{2m}\right)},$$

which is precisely the third term on the right side of the Bakhshali formula; in other words, the Bakhshali square root is $(\sqrt{n})_3$ in a sequence which can be repeated indefinitely if one wishes, the ith term obtained by linearising the error in the $(i-1)$th iterate. There are several ways of exhibiting the recursive structure of the infinitely iterated formula (which Bakhshali does not have; it stops at $(\sqrt{n})_3$); one version is given in [BM-H].

Sūtra Q2 describes the two successive correction terms exactly as they emerge in the above purported derivation. Moreover, the phrase used for the second correction is *śuddhi kṛta* meaning 'purified' which, in the context, is synonymous with the 'refined' (*saṃskṛta*) of the later terminology for (iterated) corrections. Taken together, these two facts suggest that the original rationale for the Bakhshali formula cannot have been very different from what is outlined above. And it reinforces the case for the *Śulbasūtra* $\sqrt{2}$ to have been derived by similar iterative methods, perhaps in a geometrical avatar: the first two examples, themselves separated by more than a thousand years, of a technique which eventually became very versatile and powerful. The conclusion as far as dating is concerned is that the Bakhshali square root cannot be considered to be a signature of relative lateness. And, if the suggested affiliation between the *Śulbasūtra* and Bakhshali square roots is true, it becomes no longer necessary to seek its inspiration in the very similar algorithm of Heron of Alexandria (1st century CE) as has sometimes been proposed.

It is to be noted that the Bakshali formula works just as effectively if m^2 is chosen to be greater than n – in fact, as we saw in connection with the *Śulbasūtra* $\sqrt{2}$, the method is not very sensitive to the choice of m (within reason). Another incidental remark is that the iteration of the Bakhshali 'purification' does not lead to the binomial series for $(m^2 + r)^{1/2}$ though the first three terms are the same. In fact the binomial expansion for powers which are not positive integral was unknown to Indian mathematicians.

Returning to the question of dates, the internal mathematical evidence can be summarised as indicating a date for the manuscript between approximately 300 CE and 500 CE, the lower limit with some confidence based on its fully developed decimal notation. The upper limit depends on our judgement of the level of its sophistication in relation to *Āryabhaṭīya* and, to that extent, is less certain. Perhaps as good a reason for not dating it later, in fact not later

than about 450 CE, comes from political history and geography. Gandhara and its neighbourhood were areas in which high scholarship and prosperous commerce flourished for many centuries before being reduced to ashes and rubble by the invading Hunas. That happened about 450 CE or a little earlier. The northwest ceased being a part of intellectual India after that; no Panini emerged, no Patanjali or Pingala or Vasubandhu, ever again. If Bakhshali is the village where the manuscript was actually composed (and buried), who in that post-Huna wilderness would have been there to write it and for whom would he have written it?

Though I have not mentioned them individually, many scholars have in the past come to very similar conclusions about the date, for a variety of reasons, not always the same as mine. But there are also quite a few who would prefer to place the manuscript substantially later, based primarily on non-mathematical considerations such as language, script, the physical material (birch bark), etc. Contrary to occasional claims, the use of birch bark for writing is very old in the northwest frontier; very well known is a Buddhist canonical text of the 1st century CE in Kharoshti script, found rolled up in a clay pot, presumably buried, on the western border of Gandhara. As for the Bakhshali's language (a variant of Sanskrit) and script (a variant of what later became Devanagari), scholars have disagreed quite dramatically about the period in which they were prevalent, ranging from the 2nd century CE to the 10th or later. The only certainty is that that the original Brahmi script had already begun to branch off into the several sub-scripts of the Nagari family into which it evolved, which too leads to an absolute lower limit date of roughly the 4th century. The debate is unlikely to be settled even if the remains of the manuscript are one day subjected to scientific dating techniques: what the Bodleian has may very well be a copy of an earlier original. Which, of course, will not matter if its date turns out to be very early.

7

The Mathematics of the *Gaṇitapāda*

7.1 General Survey

If the *Āryabhaṭīya* is a difficult book to read, the *Gaṇitapāda* is the most difficult of its three main parts or chapters. The economy of expression of the *sūtra* style that limits a proper understanding of the material to the already initiated becomes even more frustrating when it comes to the mathematics because we know, as did the commentators, that the topics are capable of a precise formulation, to be accepted or rejected – there *are* a few wrong statements in the *Gaṇitapāda* – on logical and mathematical grounds alone. It is quite common to find in it constructions and computations incompletely specified, essential terminolgy omitted and mathematical objects defined only partially or, at times, not at all. These obstacles to a clear reading were largely overcome over time by the mathematicians who followed him. But there was a second barrier, that of coming to terms with the novelty of concepts and methods, more especially of absorbing Aryabhata's point of view regarding the geometry of the circle in its infinitesimal aspects. The challenge of reading Aryabhata's mind, and the hope that the challenge can be met through imagination and careful study, surely played a part in the fascination that the *Gaṇitapāda* held for succeeding generations of outstandingly good mathematicians, from Bhaskara I to Nilakantha; of the very large number of commentaries produced, many in fact confined their attention to just this one chapter. Nilakantha's *Āryabhaṭīyabhāṣya* (which covers all three chapters) is especially perceptive – a stimulating combination of explanation, proof, insight and conjecture with bits of history thrown in – and it is to him that we owe the elucidation of some of the more opaque passages as well as the deeper. After all it was his great-grand guru Madhava who had finally found the key to Aryabhata's mind and the feeling is inescapable that, in this work completed towards the end of his life, the wise old man was engaged

P. P. Divakaran, *The Mathematics of India*, Sources and Studies in the History of Mathematics and Physical Sciences, https://doi.org/10.1007/978-981-13-1774-3_7

in a reaffirmation of Aryabhata's originality in the brilliant light thrown on it by Madhava's own genius.

The present section is meant to provide a general picture of what the *Gaṇitapāda* contains and to get the more routine matters out of the way before coming to grips with what is truly original and profound in it. Like many other texts, the *Āryabhaṭīya* starts with the reminder that multiplication by 10 promotes a number to the next higher place while listing the names of the powers of 10 up to 10^9, which just about covers the biggest number that occurs in the book, the number of days in a *yuga* (but does not cover the number of days in the orthodox *kalpa* of four *yuga*s – the *kalpa* is mentioned in the *Gītikā* but Aryabhata has no astronomical use for it). Unlike some other authors, he does not spend time on the basic arithmetical operations but goes straight to the geometrical meaning of squares and cubes as areas and volumes. That is almost certainly meant as preparation for the geometrical interpretation of some of the arithmetical results involving squares and cubes of natural numbers that are to come later in the chapter. Such an interpretation is not given in the text, but proofs of these results given by later authors are critically dependent on geometry.

Next are algorithms for extracting square roots and cube roots. The verses describing these procedures are among the more sketchy and incomplete in the book. But since they eventually became the basis for the standard methods of root extraction – Aryabhata's is the first documented method in India for the decimally written roots of a decimal number n and, as such, very different in spirit from the more algebraic and base-independent Bakhshali formula – the commentators seem to be agreed on the steps they are supposed to convey. Even so, they are generally silent on the essential variations in the steps that are required to accommodate special cases as determined by the decimal entries that make up the number n. The method is an adaptation of long division and the variations come about on account of the carry over rules. For the square root (at a given level of accuracy) for instance, given n as a 'polynomial in 10': $n = \sum_{i=1}^{k} n_i 10^i$, we are required to find another such 'polynomial' (after multiplying n by a sufficiently high even power of 10, to avoid working with fractions) $m = \sum_{i=1}^{l} m_i 10^i$, such that $(\sum_{i=1}^{l} m_i 10^i)^2 = \sum_{i,j=1}^{l} m_i m_j 10^{i+j} = \sum_{i=1}^{k} n_i 10^i$. The need for care arises from the fact that $m_i m_j$ may be greater than 10, causing an overflow into the coefficient of 10^{i+j+1}. The highest place l in the square root, in particular, will depend on the decimal entries of n, which is something Aryabhata does take account of, but there are also changes required in the intermediate steps which remain unmentioned. Worked examples can be found in several books.[1]

[1]Commentators including modern ones often pass over the details of these unspecified exceptional cases ([AB-C], [AB-S]; see also [DS]), limiting themselves to working out examples in which such complications do not arise. For a numerical example in which what Aryabhata does not say matters, see [Pl]. A general prescription for determining the decimal entries m_i in terms of the entries of $\sqrt{n}$, without resorting to trial-and-error, can of course be devised, but it is bound to be unwieldy.

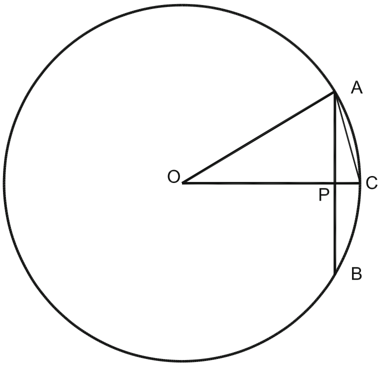

Figure 7.1: The geometry of the 12-gon from the hexagon

Similar considerations hold for the cube root.

The square root has another point of interest. In the preparation of his table of sines, Aryabhata needed the value of π which he gives as 62,832/20,000 (= 3.1416). He does not, of course, say how he arrived at the value but later authors have implied that he probably computed the perimeters of regular polygons approximating the circle by repeated use of the diagonal theorem. Thus, start with the regular hexagon inscribed in a circle of unit radius. Its side (AB in Figure 7.1) is 1, which is a result stated in the *Āryabhaṭīya* (*Gaṇita* 9). The hexagon is turned into the regular 12-gon by joining the midpoint of each arc to the neighbouring vertices (C to A and B in the figure). Then we have $AC^2 = AP^2 + CP^2 = 1/4 + (1 - OP)^2$, with $OP^2 = 1 - AP^2 = 1 - 1/4$. So the side of the 12-gon is given by

$$AC^2 = \frac{1}{4} + \left(1 - \sqrt{1 - 1/4}\right)^2 .$$

The perimeter of the 12-gon thus requires two square roots to be extracted. The doubling can be repeated, the geometry being the same for each doubling; if a_k is the side after k doublings, we have

$$a_{k+1}^2 = \frac{a_k^2}{4} + \left(1 - \sqrt{1 - a_k^2/4}\right)^2 .$$

According to the commentator Ganesha (Gaṇeśa Daivajña, 16th-17th century), 6 doublings (a 384-gon) will do to get to Aryabhata's value for π; that makes for twelve square roots to be computed which is not in itself such a challenge, but it requires the square roots in each step to be worked out to an accuracy better than five significant figures.

The reason why this point is relevant is that, assuming that he did use the polygonal approximation and did work out the square root by his own method, Aryabhata obviously could not have meant the n of the square root

algorithm to be a perfect square as is often the case in the examples worked out by commentators. But neither does he indicate how the process is to be continued below the place of ones and terminated at a desired level of precision. It is actually straightforward to carry on until the required accuracy is attained – having first multiplied n by a large enough even power of 10, as noted above – but effectively the algorithm, being designed for base 10, must deal with decimal fractions. It is a curious fact that, having mastered the decimal representation for positive (and, eventually, negative) integers, Indian mathematics never used base 10 systematically for writing fractions. Aryabhata is no exception to this as his number for π shows; so how exactly did he compute the approximate square roots to the necessary accuracy?

Nilakantha's exposition in his *Āryabhaṭīyabhāṣya* of Aryabhata's square root algorithm adds a very interesting footnote to the question of roots of non-square numbers. He too, like others before him, avoids negative powers of 10 by the expedient of multiplying n by larger and larger powers of 10 for any desired accuracy, but takes a big conceptual step forward by stating that the process is interminable: no matter how big the power of 10 chosen as the multiplier, the square root will always be approximate. That is not quite the same as saying that $\sqrt{n}$ is irrational but thinking along such lines surely influenced his claim in the same book that π is an irrational number. Nilakantha's ideas on irrationality provide a fascinating glimpse into how evolved the Nila school's approach to numbers and their logical characterisation had become, a topic that will reemerge from time to time in this book.

The polygonal approximation to the circle was a very old subject by the 5th century CE, its first use going back to the mature Harappan culture (Chapter 3.2). Before the beginning of the common era, it had been the starting point of Archimedes's method of exhaustion for finding areas and volumes of regular curvilinear geometric objects and, historically, it is not impossible that a good value of π came to India from Alexandria along with the essential elements of the new astronomy, the best candidate being Ptolemy's $377/120 = 3.1417$. But, apart from the deviation in the last decimal place, there are other circumstances which go against that hypothesis. The *Pañcasiddhāntikā* is unaware of this accurate value and is content with $\pi = \sqrt{10}$ and that too in the (post-Aryabhata) independent Chaper IV; Ptolemy's π was certainly not part of the common store of knowledge in India. More generally, knowledge of Greek geometry was minimal or absent; among many other indications, Aryabhata was ignorant of Archimedes' correct formula for the volume of the sphere, his own being wrong. On the other hand, in the early centuries of the common era there was a tremendous amount of activity in China[2] on the use of polygons with a large number of sides for determining π. The presence of Chinese scholars and pilgrims in fair numbers in India in the centuries before (and after) Aryabhata is as well established as that of Greeks earlier.

[2]Joseph W. Dauben, "Chinese Mathematics" in [MEMCII]

In short, apart from Ganesha's word, we have no basis for a good guess at where Aryabhata's value of π came from.

Verses 6 to 18 are concerned with geometry. The core of this part are the verses 10 to 12 which, without any fanfare, bring in trigonometric functions as a tool in the investigation of the geometry of the circle and hence mark the founding of a new discipline; sections 3, 4 and 5 of the present chapter are devoted to a detailed discussion of what these verses contain. These three verses are preceded by two preparatory steps: i) the area of the circle (1st line of verse 7) whose importance lies in the probable way (if we go by the writings of the Nila school) it was derived, by dividing the circumference into a very large number of arc segments (the first intimation of an infinitesimal mode of thought, see section 5 below for more on this), and ii) the introduction of the regular hexagon (2nd line of verse 9), important as the starting point of the computation of π as well as of the sine table.

The rest of the geometry is no more than a small collection of miscellaneous results, some without any obvious astronomical application or mutual thematic connection. Two of them are in fact wrong. The first is a formula for the volume of a tetrahedron (a "six-edged" solid) and it is wrong even for a regular tetrahedron. More interesting is the volume of the sphere, of radius r, which is stated to be the product of the area of the circular face of the hemisphere and its own square root, $V = \pi r^2 \sqrt{\pi r^2}$, and it is claimed to be exact (*niravaśeṣam*, with nothing left over). The culprit seems to have been a false analogy: that, once the area equivalence of a circle and a square is established, it will extend to the corresponding volume equivalence. The surprise is that Aryabhata overlooked such a gross overestimate, evident both geometrically and numerically. Especially, if he had followed the method attributed to him (by the writers of the Nila school) for the correct area of the circle as that of a rectangle of sides πr and r, it is difficult to see how the error could have been made. Homer nodded?

It was not until Bhaskara II in the middle of the 12th century that the correct volume of the sphere was obtained and it is noteworthy that the method he used is what is nowadays called finite integration based, ironically, on Aryabhata's own sine table. For a fully satisfactory treatment of the problem, however, we have to look to the Nila mathematicians who saw it as part of their calculus project, turning Bhaskara II's semi-numerical ('computational') approach into a purely analytical one. By a similar method Bhaskara II also computed correctly the area of the surface of the sphere, a topic Aryabhata and others following him had not touched. It appears that, for a long time, Indian geometry had a conceptual difficulty in dealing with curved surfaces, those that cannot be embedded in the Euclidean plane, as though it could not think beyond the natural association of areas with the plane and volumes with space; Bhaskara II's painstaking explanation of what is even meant by the area of a sphere – measure the area of a flexible surface, a net, that wraps around the sphere, after flattening it – is revealing. The evolution from the planar areas of the *Śulbasūtra* to surfaces in 3-dimensional space took a very long time.

Two verses, 17 and 18, are devoted to pre-trigonometric circle geometry. Verse 17 starts with the statement of the theorem of the diagonal, in terminology (*bhuja, koṭi, karṇa*: base, altitude and diagonal) that is more natural when applied to a rectangle, its original formulation in the *Śulbasūtra*, than to a right triangle. (It is puzzling, one of many puzzling things in the *Āryabhaṭīya*, to see the essential ingredient in the definition of the sine and the cosine coming after the sine table has already been constructed; its natural place would have been just before the value of π.) This is followed by the result: if a chord is intersected by a diameter orthogonally, then the product of the two segments of the diameter (the 'arrows', *śara*) equals the square of half the chord. (Trigonometrically, i.e., once the theorem of the diagonal is brought in, it becomes the identity $(1-\cos\theta)(1+\cos\theta) = \sin^2\theta$). Verse 18 applies this theorem to characterise how the common chord of two intersecting circles divides the line through the two centres. The proof given by later writers of these very Euclidean theorems are simple applications of similar right triangles and is in fact an adjunct of one proof of the theorem of the diagonal (see Chapter 2.2); their historical interest really is that the orthogonality properties involved are not just pre-trigonometric but pre-Pythagorean and may well have been known to the Indus town planners (see Chapter 3.3).

The last third of the *Gaṇitapāda* is concerned almost entirely with questions of an arithmetical-algebraic nature. Three main themes can be discerned: 1. Series of several different types; 2. The rule of three and systems of linear equations, and the quadratic equation in one unknown; and 3. Linear inhomogeneous Diophantine (or indeterminate) equations in two unknowns. In the rest of this section I review Aryabhata's results under themes 1 and 2; theme 3 became the seed from which grew an extensive theory of linear and quadratic Diophantine equations and for that reason will be treated separately and in some detail in the next section.

The treatment of series (*Gaṇita* 19, 21 and 22) begins with the main properties of the general arithmetic progression with initial term a, common difference d and number of terms n: partial and total sums (surprising since they are the same problem), the mean of n terms in terms of a, d and n, and as the mean of the first and the last terms; interestingly, the sum (the formula is identical with the one in Bakhshali) is computed as the product of the mean and the number of terms, as though the mean is what was obtained first. The specialisation to the sum of the first n natural numbers, $S_n^1 = S_1(n) = n(n+1)/2$, is not given, we do not know why, but the sum of sums (called the 'volume of the pile', *citighana*), $S_2(n) = \sum_{i=1}^{n} S_1(i)$, is given in two different explicit expressions:

$$S_2(n) = \frac{n(n+1)(n+2)}{6} = \frac{(n+1)^3 - (n+1)}{6}.$$

This is followed by formulae for the sums of squares and cubes of natural numbers in the forms

$$S_n^2 = \sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6},$$

$$S_n^3 = \sum_{i=1}^{n} i^3 = (S_n^1)^2.$$

All three of the summation formulae have important historical links. First, the sum of sums is also the binomial coefficient ${}^{n+2}C_3$ and, in that combinatorial avatar, is a special case of the answer to Pingala's question Q5, the number of metres with $n+2$ syllables of which 3 are **g** or **l** (see Chapter 5.5). Almost none of the commentators seems to have made the connection, especially surprising in those who came after Halayudha's explication of the combinatorics of Pingala in the 10th century CE; it is as if prosodists and astronomers did not talk to one another. This particular type of sums of sums, called *saṃkalita-saṃkalitam* or equivalent terms in the Nila texts and defined recursively by $S_k(n) = \sum_{i=1}^{n} S_{k-1}(i)$ for general n and k, were destined to play an absolutely crucial role in the development of the sine and cosine series by Madhava. By then Narayana in mid-14th century had found the connection between sums of sums and combinatorics and had the general formula for $S_k(n)$. Nevertheless, no Nila author gives its proof, with one of them, Shankara (Śaṅkara Vāriyar), going to the extent of saying that it is too difficult for those with limited intelligence to understand.

The reason for this disconnect has to do with the way Aryabhata himself very likely arrived at his results on series. In a return to the *Śulbasūtra*'s cut-and-paste methods in area-equivalence problems, but now elevated to one dimension higher so as to accommodate cubic terms, authors who have provided rationales for them generally follow what may be called a building-block method in which suitably shaped and sized bricks are put together, with a great deal of ingenuity and 3-dimensional imagination, to build volume-equivalent solid bodies – geometric arithmetic/algebra again. In the absence of any direct evidence, and with faith in the robustness of the oral tradition of teaching, we have little choice but to accept that the commentaries reflect what came down to them from the master over the generations. Nilakantha's *Āryabhaṭīyabhāṣya* describes them in detail (see [SA] for a careful description of these 'architectural' constructions). There is support for the view that Aryabhata used just such geometric methods in the term employed, 'volume of the pile' rather than 'sum of sums', for $S_2(n)$, as well as in the reminder of the association of squares and cubes with areas and volumes in *Gaṇita* 3 and in the fact that the sum of cubes is expressed as an area squared rather than as an expression of explicit degree 4 in n which would have called upon a 4-dimensional imagination.

The generalisation of the formulae for the sums of squares and cubes to arbitrary powers was also to play a central role in the calculus work of the Nila school. Solid geometry was of no help here. What was required is the behaviour of the general power sum $S_n^k = \sum_{i=1}^{n} i^k$ in the limit of large n (in the integration of x^k) and *Yuktibhāṣā* does not even bother to cite Aryabhata's exact results for

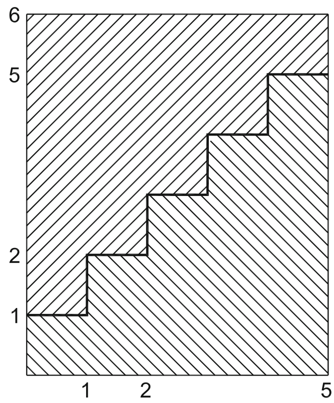

Figure 7.2: The sum of natural numbers as an area

$k = 2$ and 3. Instead of looking to geometry, it uses mathematical induction, described more or less formally in strictly algebraic language, to relate the asymptotic behaviour of S^n_{k+1} for large n to that of S^n_k for all k. Geometry is used only in the first step of the induction, to evaluate S^n_1 as the quadrature of the 'staircase function', by a construction which is planar and which almost certainly predates Aryabhata: represent S^n_1 by the area of its histogram and fit two copies of it together as shown in the figure.

On the other hand, *Yuktibhāṣā* does not say how it evaluated – exactly, not asymptotically – the general sum of sums $S_k(n)$ (no Nila text does), but it is almost certain that it was also done inductively. Thus did the demands of the new mathematics of Madhava force the abandonment of the long association of geometry with arithmetic and algebra and give birth to a novel proof-method, that of mathematical induction.

The basic algebra of linear and quadratic equations (theme 2 in the list above) has about it the air of a recapitulation of known facts and so can be summarised quickly; some of the problems are in fact to be found in the Bakhshali manuscript. It begins with the identity $xy = [(x+y)^2 - (x-y)^2]/4$, whose geometric counterpart in terms of areas is a throwback to the *Śulbasūtra*, and then goes on to the determination of x and y when $x - y$ and xy are given. One can do this by substituting and solving the resulting quadratic equation or by using the area identity[3] $(x-y)^2 + 4xy = (x+y)^2$. This is followed by a problem in discretely compounded interest (given the principal and the total returned after a given number of periods when compounded at the end of

[3]Problems of this type in which one is given the values of a subset of a set of elementary functions of x and y and is required to find the values of the whole set, with or without solving for x and y, never seem to have lost their appeal. As late as in the mid-16th century, the highly respected but relatively less well-known Nila mathematician Chitrabhanu (Citrabhānu) took them up for cubic polynomials and solved them both geometrically (by dissecting and refitting appropriate solid bodies) and algebraically ([SA]).

the first period, what is the interest per period?) presumably as an exercise in solving quadratic equations.

Next (verse 26) is the principle known as the rule of three (*trairāśikam*): if four quantities are in proportion, $a : b = c : d$, three of them determine the fourth, $a = bc/d$ along with similar equations for b, c and d. Throughout the course of Indian mathematics, the rule of three is the preferred way of formulating linear (or inverse linear) relationships among quantities: if (x_1, y_1) and (x_2, y_2) satisfy the equation $y = ax$, then $y_2 : x_2 = y_1 : x_1 (= a : 1)$ and if the equation is $y = ax + b$, then $(y_2 - y_1)/(x_2 - x_1) = a$. The rule has several straightforward variants (inverse linearity, linear dependence on more than one variable, etc.) and plays a primordially important role not only in linear problems but also as the default first approximation in non-linear problems which are sought to be solved by the method of iterated refining. It has a geometrical version that starts by approximating an arc as a straight line (to which *trairāśikam* can be applied) 'locally' and refining it successively. This is the guiding philosophy which led Aryabhata to his sine difference formula and Madhava, in due course, to the rectification of the circle (the infinite series for π) and the power series for sine and cosine.

Aryabhata's statement of the principle is numeric/algebraic and comes ahead of his treatment of (systems of) linear equations. The issues addressed are elementary: i) how to set up and solve a simple linear equation from given data, with two illustrative examples (one concerning uniform relative motion of two bodies in a straight line, presumably with astronomical applications in mind); ii) a general recipe for actually solving a linear problem, in which the equation results from subjecting the unknown to the four basic operations with known numbers (constants) in arbitrary order, by inverting them in the right sequence; and iii) the solution of a particular system of equations

$$-x_i + \sum_{j=1}^{n} x_j = a_i,$$

$i = 1, \cdots, n$, by adding them up:

$$(n-1) \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} a_i.$$

The choice of these topics is not arbitrary. They come into play in the algebraic component of the sine difference formula; in particular, the system of linear equations satisfied by the tabulated sine differences is precisely of the type considered in iii).

A notable feature of the entire *Gaṇitapāda* is that the approach is very theoretical and general. Statements are made about a general problem and its solution rather than about special cases and (numerical) examples. Indeed, apart from the value of π, there are no numbers in the chapter (the numerical sine differences are in *Gītikā*); even the two concrete illustrations of the solution

of the quadratic equation (the number of terms in an arithmetic series and the interest rate problem) are formulated algebraically. Though he does not set down the rules governing basic algebraic operations as Brahmagupta did soon after (his extreme form of the *sūtra* style leaves no room for such pedagogic indulgences), it is clear that he was completely at home in the setting of an algebraic mode of thinking.

7.2 The Linear Diophantine Equation – *kuṭṭaka*

Technically the most challenging mathematics in the *Gaṇitapāda* is in the last two verses, 32 and 33. The problem they deal with can be stated in several equivalent ways. The most natural, in the light of later developments, is as a Diophantine problem: find all positive integral values of x and y satisfying the equation

$$by - ax = c$$

in which the coefficients a and b are positive integers and c is an integer, positive or negative. Partly because the question was so simply and clearly posed in a framework in which Indian mathematics had long been comfortable – the arithmetic of integers – and required no mental readjustment to strange new concepts, it attracted an enormous amount of activity in the following centuries, from Bhaskara I and Brahmagupta to Bhaskara II and Narayana (mid-14th century). The commentators added several refinements to the problem and brought in new insights but Aryabhata's original method of solution, called *kuṭṭaka* or *kuṭṭākāra* by his followers and commonly translated as 'pulverisation', survived all these elaborations. The method relies on a technique of reduction in stages, using the Euclidean algorithm, to a sequence of equivalent equations with smaller and smaller coefficients, the last of which can be solved trivially ('by inspection'). Whether it should be considered as belonging to the discipline of arithmetic or algebra (as Indian mathematicians did) or as the initiation of a new discipline, number theory, is a matter of taste. In any case, it is the precursor of what was later called Fermat's method of descent in Europe (whose unsung prototype is the decimal representation of a number through repeated division by 10, barely heard of in the Europe of Fermat), used here to find the solution of a number-theoretic problem rather than to prove its non-existence as in Fermat's first use of it. To go back to a point that concerned us earlier, there is nothing in the Bakhshali manuscript that comes near it in originality and sophistication.

Apart from its mathematical interest, the problem is of paramount importance for Aryabhata's system of the world as defined by the periodicities of celestial bodies and, especially, for his idea of the *yuga*. (That is the other reason for the popularity of the subject among later astronomers). His own formulation of the problem is not as a linear indeterminate equation *per se*, but in an equivalent form, often called the Chinese remainder problem: find positive integers N such that, for given positive integers a, b, $\rho < a$ and $\sigma < b$, ρ and

σ are respectively the remainders when N is divided by a and b; equivalently, solve the equations

$$N = ax + \rho, \qquad N = by + \sigma,$$

expressing the congruences $N \equiv \rho \bmod a$ and $N \equiv \sigma \bmod b$, in positive integral x and y. On eliminating N this becomes equivalent in its turn to asking for the solution of the Diophantine equation $by - ax = \rho - \sigma$; conversely, a solution (x, y) of this problem and a choice of σ satisfying $\sigma < b$ and $c + \sigma < a$ give directly a solution of the remainder problem.

In its remainder form the problem arises very naturally in time-keeping with two different clocks measuring time in the same unit. In astronomical applications, the clocks may be two 'planets', say the sun and the moon, and the unit of time in terms of which their periods are integral is determined by the accuracy demanded, call it a day (it is not important from the theoretical viewpoint how a day is defined). Knowing the periods a and b and knowing also that ρ and σ days respectively have elapsed in their current periods, one wishes to know how many days have passed after the day (the epoch, $t = 0$) on which the initial days of their periods coincided. In principle, again, the choice of the epoch is a matter of convention but it is practically convenient to choose it as marking some special configuration of the planets, when they are 'in conjunction'. (One of the issues Brahmagupta had with the *Āryabhaṭīya* was that it counted the epoch from the dawn of a particular day in Lanka, determined by counting back, while he preferred the midnight). Whatever the convention used, the answer to the question is the smallest N satisfying the conditions above and it does not depend on how many full periods have elapsed; once N is found, the number of individual completed periods (x and y in the remainder problem) is also known.

In fact one knows more. If m is a common multiple of a and b, then $N + m$ is also a solution of the remainder problem; if N_0 is the smallest solution and m is taken to be the least common multiple $M(a, b)$ of a and b, the general solution is $N_0 + kM(a, b)$ for k any non-negative integer. But this general solution also solves the remainder problem with the common period $M(a, b)$ as the divisor and N_0 as the remainder, and so can be made the starting point for extending the method to more than two divisors: to find $N(1, 2, 3)$ such that it leaves ρ_i as the remainder when divided by a_i ($i = 1, 2, 3$), first find $N(1, 2)$ solving the problem for $i = 1, 2$, then solve the congruences $N(1, 2, 3) \equiv N(1, 2) \bmod M(a_1, a_2)$ and $N(1, 2, 3) \equiv \rho_3 \bmod a_3$. (In the special case of vanishing remainders, $\rho_i = 0$, the procedure reduces to the useful formula $M(a_1, a_2, a_3) = M(M(a_1, a_2), a_3)$). And so on, recursively.

Stanzas 32 and 33 do not seem to have a reference to the extension to more than two divisors, at least not according to most modern readings and translations. But Bhaskara I's *Āryabhaṭīyabhāṣya* does indicate, with worked examples, that such an extension is meant to be understood and there is circumstantial support for that position in Aryabhata's idea of the *yuga* and in the way it is made use of. The computation of its duration does not need the

remainder problem to be solved; theoretically it is just the least common multiple of all the planetary periods. But the determination of the epoch from the number of days elapsed to the current day in each planet's cycle (the remainders) does require the solution of as many congruences as there are planets. It is to be noted that Aryabhata informs us of his year of birth in precisely such terms: "When 3,600 years and 3 quarter-*yuga*s had elapsed, 23 years had passed after my birth." (*Kālakriyā* 10).

Later mathematicians who wrote on *kuṭṭaka* often rephrased Aryabhata's original problem in other equivalent ways. A popular formulation was: given positive integers a, b, c, find all integral x such that $(ax+c)/b$ is integral; others can be obtained from this by transposition, cross multiplication and division, etc. These variant formulations played a role in the terminology; for instance, b can be a multiplier in one formulation and a divisor in another; c is variously called an additive or subtractive constant or an interpolator. Also, the cases c positive or negative are generally considered distinct though the difference is only an artefact of notation, a matter of interchanging the two unknowns x and y. Indian mathematicians tended to the literal in their terminology, based as it was on a faithful functional description – 'multiplier' and 'multiplied' for example for the two factors in a product – further aggravated by the absence of a symbolic notation. In the voluminous literature on *kuṭṭaka*, although the inherent unity of the various formulations was fully understood, this can be a source of confusion for the modern reader. Those who would like to see these variations treated individually will find it done exhaustively and with care in the book of Datta and Singh ([DS], Chapter III (which is all of the 2nd volume), sections 13-15).

As a result of all this activity, the method of *kuṭṭaka* had evolved into a complete and coherent theory of the linear Diophantine equation by the time Bhaskara II wrote his *Bījagaṇita*. The following account is meant to give an idea of the theory as seen from that vantage point.[4] Its algorithmic implemention is described in a number of books and articles.

Euclidean division and the elementary properties of continued fractions form the main technical component of the theory. That is natural and the reason is clear already in the trivial case in which the remainders ρ and σ are zero. N is then a multiple of the least common multiple $M(a, b)$. The most economical practical method of computing $M(a, b)$ is to first find the greatest common divisor $D(a, b)$ and then to use the relation $M(a, b)D(a, b) = ab$. And the most efficient algorithm for $D(a, b)$ is Euclidean division. Since it plays an

[4]A synthetic account of the theory of the equation as it finally came to be in India is available in the article of A. K. Dutta, "Kuṭṭaka, Bhāvanā and Cakravāla" in [SHIM]. The article emphasises the unifying thread connecting the contributions of individual mathematicians to the creation of the theory and so is valuable both historically and from the purely mathematical angle. Dutta also draws parallels between the Indian work and later European developments in the subject. Another excellent reference for an overall picture of the theory and practice of *kuṭṭaka*, with the emphasis on continued fraction methods, is the Appendix (in English) titled 'Kuṭṭākāram' to the pioneering annotated Malayalam edition of *Yuktibhāṣā* of Tampuran and Ayyar ([YB-TA]).

essential role in more than one aspect of the general *kuṭṭaka*, I recapitulate it briefly.

Suppose $b < a$. The Euclidean algorithm is defined by the sequence of steps $a = n_0 b + r_1$, $b = n_1 r_1 + r_2$, $r_1 = n_2 r_2 + r_3$ and so on, with $r_{i+1} < r_i$. (If we denote a by r_{-1} and b by r_0, we can write this set of equations in a unified way as $r_i = n_{i+1} r_{i+1} + r_{i+2}$). The sequence ends when r_{k+1} vanishes for some finite k, i.e., $r_{k-1} = n_k r_k$, which it will since a and b are positive integers and $0 < r_{i+1} < r_i$; then $D(a,b) = r_k$. The proof is simply the observation that since r_k divides r_{k-1}, it divides $r_{k-2} = n_{k-1} r_{k-1} + r_k$ and so on. Moreover, working backwards, we have $r_k = -n_{k-1} r_{k-1} + r_{k-2} = -n_{k-1}(-n_{k-2} r_{k-2} + r_{k-3}) + r_{k-2} = (n_{k-1} n_{k-2} + 1) r_{k-2} - n_{k-1} r_{k-3}$ and so on, ultimately expressing $D(a,b)$ as an integral linear combination of a and b.

The result that the greatest common divisor is the last non-vanishing remainder in the Euclidean algorithm is often called the Chinese remainder theorem. It was surely known to Aryabhata not only because it is the heart of the technique of *kuṭṭaka*; very likely, he determined the duration of the *yuga* as the least common multiple of the planetary periods through the nested recursive formula written down earlier, $M(a_1, a_2, a_3) = M(M(a_1, a_2), a_3)$, and relating each binary M to the corresponding D, rather than by finding the prime factors of all six of them (the period of the earth supplying the unit in which they are measured). As we saw in Chapter 6.2, the first occurrence of Euclidean division in India is in the pre-Aryabhata ('Greek') Siddhanta texts, though they used it for a different purpose, to simplify fractions with large numerators and denominators by means of convergents of their continued fraction representations. Subsequent to Aryabhata, both Bhaskara I and Brahmagupta knew about its use to find the greatest common divisor and they state explicitly that when divided by the 'last (non-vanishing) residue', r_k in our notation, a and b become relatively prime, in other words that r_k is $D(a,b)$. Whether there was a Chinese influence in the Indian interest in the technique is however problematic. The first Chinese mention of a simple remainder problem has been dated to the 4th-5th century CE, a little after the early Siddhantas (mid-2nd to mid-4th centuries, see Chapter 6.2), and it was not until the 13th that the method is clearly described (Dauben in [MEMCII], cited above). Given the intensity of India-China contacts in the intervening centuries, especially the dominant role of Indian expatriates in Chinese astronomical activity (see section 5 below) it is, on the contrary, more likely that the influence was the other way. The 'Chinese' in its apellation presumably had its origin some time after the (re)introduction of the problem to Europe by Fibonacci in the 13th century, just like the 'Arabic' in Arabic numerals.

The connection of the Euclidean algorithm to the solution of the Diophantine equation is more or less direct. Assume as before that $a > b$ and also that b does not divide a, so $D(a,b)$ is strictly less than b. Write the linear dependence of $D(a,b)$ on a and b as $D(a,b) = a\xi + b\eta$, with ξ and η integral. Then, from $a\xi + b\eta < b$, it follows that ξ and η cannot both be positive ($b > 1$). From the positivity of $D(a,b)$, $a\xi + b\eta > 0$, we also have that ξ and η cannot both be

negative. If ξ is negative and η is positive, redefine $X = -\xi$ and $Y = \eta$; in the alternate case, $X = \xi$, $Y = -\eta$. Covering both cases (as well as the case $b > a$), the pair (X, Y) is a positive integral solution of

$$bY - aX = \pm D(a, b).$$

We can now divide through by $D(a, b)$, resulting in the special linear Diophantine equation

$$BY - AX = \pm 1$$

with positive integral coefficients $A = a/D(a, b)$ and $B = b/D(a, b)$ which are relatively prime, to be solved in positive integers. All the variations of *kuṭṭaka* are effectively equivalent to reducing the general problem with $c \neq \pm 1$ to the special problem.

To begin with, the general linear Diophantine equation has no solution unless $D(a, b)$ divides also c; otherwise, on dividing by $D(a, b) \neq 1$, the left side will still be an integer while the right side becomes a proper fraction. So it will be understood from now on that a and b are relatively prime; all writers beginning with Bhaskara I state that the first step in doing *kuṭṭaka* is to make the equation irreducible by dividing by $D(a, b)$. After that, the approach varies in detail, but not in essence, from author to author. The account below is based on that of the appropriately named Aryabhata II (*Mahāsiddhānta*, 10th century; there is no need to attach I to the name of the first Aryabhata as his later namesake only makes this one cameo appearance in this book). The reasons for this preference are twofold: Aryabhata II is relatively easy to follow since he explains all steps systematically and clearly (see [DS] for a translation of the relevant passage) and his approach is very structural and so more in line with the tastes of the modern reader. Apart of course from the notation, the only liberty I have taken is to streamline the presentation slightly by reordering some of the steps. Also to be kept in mind is that other traditional commentators do not always adopt such an abstract approach and generally illustrate the issues involved with numerical examples.

After making the equation irreducible, the next step is to note that if (x, y) and (x', y') are any two solutions, then $(y' - y)/(x' - x) = a/b$ and so they are related by $x' = x + mb$ and $y' = y + ma$ for integral m. Since a/b is positive, $x' - x$ and $y' - y$ have the same sign and therefore there is a solution (x_0, y_0), the minimal solution, such that i) for any other solution (x, y), $x > x_0$ and $y > y_0$ and ii) it generates all solutions: $x = x_0 + mb$ and $y = y_0 + ma$ as m is varied over all positive integers.

Next, given a solution (x, y), divide x by b and y by a:

$$x = kb + r, \qquad y = la + s,$$

with $0 < r < b$ and $0 < s < a$. If $k = l = 0$, we get a formal solution as a function of either of the two residues r or s:

$$x_0 = r, \qquad y_0 = s = \frac{ar + c}{b}$$

or, equivalently,

$$x_0 = \frac{bs - c}{a}, \qquad y_0 = s.$$

This solution is in fact minimal, independently of the sign of c, since x cannot be less than r and remain positive (and similarly y; as we saw, one condition implies the other). Aryabhata II proceeds somewhat more elaborately (he is interested in more than just the shortest route to the solution) and takes up the cases of positive and negative c separately. The condition $c > 0$ becomes, on substitution of the general form of the ansatz, the inequality $(l-k)ab+bs-ar > 0$; replacing r by its minimum value 0 and s by its maximum value a, we get $ab(l-k+1) > 0$ or $l-k \geq 0$. Then $x_0 = r$, $y_0 = s + (l-k)a$ is a (positive integral) solution, in fact the minimal solution. If c is negative, we can repeat the argument above to conclude that $l - k \leq 0$ and that the minimal solution is $x_0 = r - (l-k)b$, $y_0 = s$. In either case, $l-k$ can be eliminated in favour of c and the result is, as expected, the expressions for x_0 and y_0 displayed at the top of this paragraph; the minimal solution (and hence all solutions) is determined by one of the residues r or s.

The final question is thus: how does one find the minimal solution? Aryabhata II, and others before and after him (the two Bhaskaras and Brahmagupta among them), note that it is enough to solve the problem for a convenient choice of c, most naturally $c = \pm 1$. If (X, Y) solves the special equation $bY - aX = \pm 1$, then (cX, cY) solves the general equation and all solutions of the general equation arise in this way. If R and S are the residues in the special equation, the corresponding residues in the general equation are cR mod b and cS mod a. (One comes to the same conclusion from the dependence of the minimal solution on c). But, as we saw earlier, the solution of the special equation is just the inversion of the sequence of Euclidean divisions used for finding the greatest common divisor $D(a, b)$, specifically the determination of the coefficients X and Y in the expression of $D(a, b)$ as a linear combination of a and b.

Aryabhata himself does not have anything to say about these theoretical points, nothing about the special case $c = \pm 1$, not even about the initial simplification of dividing by a common factor. Modern translations of *Gaṇita* 32 and 33 generally incline towards one or the other of the interpretations given by particular commentators and they are not always, literally speaking, the same. But about the contents they seek to convey, once a coherent interpretation of the bare-bones verses is agreed upon, there is no dispute: it is the Euclidean algorithm and its inverse, whose sequential steps are implemented in a tabular form, adapted from its use in finding $D(a, b)$ to the situation in which c does not divide a and b. Its 'secret of success' is that successive Euclidean divisions give rise to smaller and smaller remainders and hence replace the original equation by equivalent equations with smaller constants, justifying the name *kuṭṭaka*, first used by Bhaskara I. All the simplifications and reductions that later authors introduced had the same object, that of 'pulverising' the constants. The most interesting of them, which Aryabhata II makes much of, is the following. If $D(b, c) \neq 1$, divide the general equation by it:

$[b/D(b,c)]y - a[x/D(b,c)] = c/D(b,c)$; so it is enough to redefine x as $x/D(b,c)$, solve $[b/D(b,c)]y - ax = C/D(b,c)$ and scale up x. It is straightforward to generalise this reduction to include the possibility that a and c also have a common factor, thereby reducing the size of a as well.

7.3 The Invention of Trigonometry

The first text ever to have a word signifying the sine is the *Āryabhaṭīya*. The word is *jyārdha*, meaning half the chord; it occurs first in *Gītikā* 12 and *Gaṇita* 11 and 12, the verses concerned with the sine table (and, subsequently, at several places in the astronomical chapters) and the literal meaning makes clear that the sine was associated conceptually with the arcs and chords of a circle rather than with the angle between two decontextualised straight lines. The word is not defined in the text nor the logic behind its introduction explained. And, strictly speaking, the term is inaccurate because it associates to an arc not half its chord as the literal meaning would imply – that would not have been an interesting thing to do, neither requiring any geometric insight nor leading to new consequences – but half the chord of twice the given arc, as shown in Figure 7.3. The hyphenated 'half-chord' will be used in that meaning from now on.

The arc AB is doubled to the arc AC and the intersection of the line OB with the chord AC defines the right triangle APO. It is this triangle that is associated to the arc AB; in other words, the arc AB (or equivalently the angle θ it subtends at the centre) is not directly sought to be associated to its own chord, but defines uniquely two straight lines whose lengths are related to the radius R by the diagonal theorem: $AP = R\sin\theta$, $OP = R\cos\theta$, with $AP^2 + OP^2 = R^2$. In the Indian view and in modern language, the functions sin and cos are defined by $\sin(\text{arc}AB) = AP$ and $\cos(\text{arc}AB) = OP$; the function tan, though it is useful in shadow-geometry problems, is considered a derived concept which attains real prominence only in one context, that of Madhava's π series.

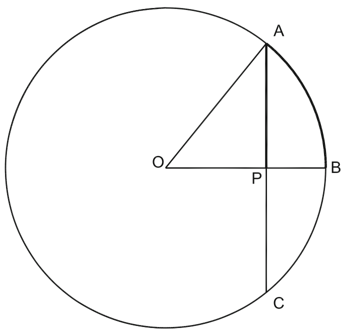

Figure 7.3: The arc and its half-chord

This simple construction by which the geometry of the circle is reduced to that of suitable right triangles can be said to mark the birth of trigonometry as an independent discipline of its own. This is not a unanimously held opinion. As in the case of other mathematical disciplines or subdisciplines, there is no universally valid position on when the geometry of the circle made the transition to trigonometry, with or without the prefix 'proto'. It would nevertheless seem reasonable to think of trigonometry, as the subject is understood today (and since a long time), as the study of trigonometric functions, whether defined implicitly and geometrically, with their periodicity properties directly rooted in the circle (as above) or explicitly and analytically (as in later work). The present section is meant to set the stage for an assessment of this position, as a summing up, in relation to its historical context, of what is trigonometric in the *Āryabhaṭīya*. Naturally and inevitably, questions of origins, influences and provenance will come up, and will be dealt with as they arise, in the rest of this section and the next two.

The founding idea of Aryabhata's trigonometry is the recognition that a diameter and a chord intersecting each other at right angles is a good pair of lines in terms of which to express the metrical properties of the circle. The 'metric' in the geometry comes from the theorem of the diagonal as it is stated geometrically in the *Śulbasūtra*. The other essential element, pairs of orthogonally intersecting chords and diameters, is equally ancient if not older, its prototype being the common chord and the line of centres of intersecting circles as used in Baudhayana's construction of the square or even in the determination of cardinal directions by the shadow method, perhaps going back to the Indus culture (Chapter 2.3 and Chapter 3.3) – in any case, definitely pre-Pythagorean (in the sense also that it does not depend on the diagonal theorem). Aryabhata himself presents these two facets sequentially in the same verse (*Gaṇita* 17) and, later, Bhaskara II brought them together by basing one of his proofs of the diagonal theorem on the right triangle defined by the diameter as the hypotenuse and a perpendicular half-chord as the corresponding altitude (see Chapter 2.2, Figure 2.6).

Given the antiquity of these ideas, why did it take so long for them to come together? Part of the answer may be that geometry got frozen into an unthinking ritualistic mould after the early *Śulbasūtra* – there is not a great deal that is geometrically original in Katyayana's *Śulbasūtra* as compared to that of Baudhayana or Apastamba, though he lived 400 or 500 years later. It would seem that all the mathematical energies had gone instead into the exploration of numbers (and the concomitant development of arithmetic) and other abstract areas like combinatorics, not to forget grammar itself. And it took the demands of the Greek-inspired new astronomy to reawaken a subject which had been dormant, but not dead, for more than a millennium.

The conceptual richness and computational simplicity thus brought about was reinforced by an elementary but astute convention. The half-chord is a length that scales linearly as the radius. Once it is recognised that a good measure of an arc, invariant with respect to changes of scale of the radius, is

the angle it subtends at the centre, a scale-independent characterisation of all circles is as having a circumference C of 360 degrees. Since the radius and the circumference also scale together, all circles can be assigned a standard radius R_0 defined by $C = 2\pi R_0$ or $R_0 = 360/2\pi$ in degrees. R_0 is a constant determined by the units in which C (and hence the radius) is measured: for instance, if C is measured in radians, $R_0 = 1$ and we may think of (the equivalence class of) all circles as being represented by the unit circle, as we do these days. The important thing is to measure both the circumference and the diameter in the same units, degrees and minutes in Indian trigonometry. Aryabhata prefers to work in minutes of arc (convenient for small angles; his smallest angle/length is 225 minutes) and so his standard radius is $R_0 = 21,600/2\pi$ in minutes. All lengths that scale as the radius, for example the half-chords, are also measured in minutes, normally thought of as a circular measure. A good value for π is therefore obligatory for doing trigonometry.

Aryabhata gives his own value of π as approximately 3.1416: *Gaṇita* 10 says that the circumference of a circle of diameter 20,000 is very close to 62,832. For this value of π, the constant R_0 is 3438 minutes correct to the nearest minute. (Some later astronomers worked to a higher degree of precision). The same constant is involved in the numerical values of half-chords, measured in minutes of arc. Modern commentators often denote the half-chord in minutes as Rsin; the use of Rsin and Rcos is strictly equivalent to working with our usual sin and cos and all trigonometric relations expressed in terms of Rsin and Rcos remain true when they are replaced by sin and cos and arcs are replaced by angles in radians. I shall follow the modern notation throughout this book. (I will also not have much use for the function versine (the length of PB in Figure 7.3) called the arrow, *śara*, in India and often preferred to the cosine).

Thoroughly familiar as we are with the basic trigonometric functions, all these lengthy explanations about measurement and units may appear superfluous and unnecessary. It is a historical fact nevertheless that Greek geometry had a problem in dealing with the issue in practice, in particular in quantitative astronomy. The difficulty had its roots in the divorce between numbers and geometry and showed itself in the fact that angles were measured in degrees and their sexagesimal fractions – not in arc lengths – while essentially arbitrary units where employed for linear dimensions. Ptolemy for instance divided the radius into 60 parts in the calculations for his table of chords, handy for sexagesimal arithmetic but a burden otherwise as it required endless manipulations with constants of conversion to get the numbers right. It is as though the fact that $C = 2\pi R$ holds only if arcs and chords are measured in the same unit was lost sight of. One might even add as a corollary that there is no useful 'metry' in trigonometry without a common, standard unit. The Ptolemaic shadow stayed over Europe for a long time, even after trigonometry became an indispensable skill for cartographers and navigators - - the radian, in all but name, began to acquire currency in numerical trigonometry only in the 18th century. It is not unusual, even today, to see statements to the effect that the main advantage

of the radian as the angular unit is that it helps to integrate geometry into anaysis. That it is, as already illustrated in Aryabhata's view of trigonometry and what became of it later in the work of the Nila school; but it also lightened the labour of the computation of sines enormously.

It is an interesting coincidence (or, perhaps, not) that, at a time when the radian had not yet been fully adopted, John Playfair, the Scottish geometer and historian of mathematics – the same Playfair who reformulated Euclid's parallel postulate – published an article (in 1798) in which he compared Aryabhata's table of half-chord differences with Ptolemy's chord table. It has an extensive passage on the role played by units in the preparation of the two tables and, by extension, in the genesis of modern trigonometry (which passage is the direct inspiration for the paragraphs above). Playfair's was the first ever learned article on Indian trigonometry – on Indian mathematics in fact, as far as I can tell – published in Europe in a European language.[5] That alone should make it compulsory reading for the historian. As it happens, it is also among the most perceptive analyses of the central themes of Aryabhatan trigonometry. Present day historians do not always seem to be aware of Playfair's careful discussion of the historical and mathematical questions at issue, especially illuminating on the problem of what consitutes or should be taken to constitute the decisive new idea separating trigonometry from Euclidean circle geometry. And, adding to its relevance, he was writing at a time (between Euler and Fourier) when trigonometry in Europe had settled into a mature discipline making rapid strides, especially in the analytical study of trigonometric functions.

Playfair's essay was not based on the *Āryabhaṭīya*. What he had to go on were the translations of two brief extracts from the *Sūryasiddhānta* (published in Calcutta a year later) by a Mr. Samuel Davis, well-known in the Calcutta circle ("Mr. Davis who first opened to the public a correct view of the astronomical computations of the Hindus" – Colebrooke [Co], Dissertation). This *Sūryasiddhānta* is not the *Saurasiddhānta* of the *Pañcasiddhāntikā*, but a text which, in most of its substance, is later, very likely the text translated and edited by Burgess ([SuSi-B]) in the middle of the 19th century. At the time of Davis and Playfair, Aryabhata was little more than a name (Davis has a few mentions of Āryabhaṭṭa, cited from other texts) in the very narrow circle around William Jones in Calcutta, the body of his work remaining beyond their knowledge. (It took another eighty years for the first printing of *Āryabhaṭīya* (the edition of Kern) to see the light of day). Knowing what we know now, there can be no doubt that the mathematics that Davis and Burgess translated, including the computation of the sine table, is to be attributed to the *Āryabhaṭīya*; no text earlier than it has any trigonometry (see Chapter 6.2 on the *Pañcasiddhāntikā*) and the trigonometric material in virtually all later texts was no more than readings and elaborations of it.

---

[5] "Observations on the Trigonometrical Tables of the Brahmins", Transactions of the Royal Society of Edinburgh, vol. **IV** (1798), p. 255. Playfair's statements on the absolute chronology of Indian trigonometry are entirely to be discounted.

The thrust of Playfair's article is on reconstructing the rationale behind Aryabhata's table and a remarkably good job – considering how little he had to go on, just the statement of the rule plus the numerical values – he makes of it. It was inevitable that he would compare it with Ptolemy's, the motivation and purpose of the two being the same, that of computing the chords corresponding to a discretisation of the circle fine enough for the precision demanded by the observational astronomy of the day.[6] But, as Playfair explains at length, the means by which they got to their objective were very different. Ptolemy did not find the geometry he needed in the extant knowledge and had to invent a new theorem which was, nevertheless, entirely Euclidean in spirit and proof: in a cyclic quadrilateral, the sum of the pairwise products of opposite sides is the product of the two diagonals. He then derived corollaries, in particular on the chords of 'sums' and 'differences' of contiguous arcs, with the help of which he could calculate the chords of multiples of 90 minutes of arc starting with a few easy standard chords (yet more work was needed to get to 30 minutes). It is common these days to present the corollaries in trigonometric notation as the addition theorems for sines: $\sin(\theta \pm \phi) = \sin\theta\cos\phi \pm \cos\theta\sin\phi$. But Ptolemy had no sines and no cosines and his statements and proofs of the addition theorems are fully classical-geometric in language and method, everything in terms of full chords; right triangles do figure in the construction but the right angle is the 'Greek right angle', the one subtended by the diameter at a point on the circumference, and not the 'Indian right angle', the intersection of the diameter with the half-chord.

The only 'pure' trigonometry to be found in the *Āryabhaṭīya* – i.e., apart from applications to quantitative astronomy – is in the table of half-chord differences (or, more loosely and dropping the scale factor R_0, the sine table as it is generally referred to), almost as though it was invented for the express purpose of computing sines. Actually, despite Aryabhata's term *jyārdha* (equivalently, *ardhajyā*) rather than the correct *ardhajyāntara*, it is a table of sine differences, the differences of consecutive sines in arc steps of $\epsilon = \pi/48$ (=225 minutes) in the first quadrant $(0, \pi/2)$:

$$\delta s_m = s_m - s_{m-1} \ \ (m = 1, \cdots, 24), \qquad s_m = \sin m\epsilon.$$

Each one of the characteristic ways in which Aryabhata's table differs from Ptolemy's (aside of course from the use of sines) – i) it is a table of sine differences, ii) it stops at $\theta = \pi/2$ and iii) the step size is $\pi/48$ – carries historical significance. As noted in connection with Varahamihira's sine table (Chapter

[6]It is arguable whether the available precision of observational astronomy justified the enormous computational efforts of Ptolemy (three sexagesimal places for chords in steps of half a degree). On the Indian side, it may come as a surprise to some that Aryabhata, and Indian astronomy in general, was definitely interested in observation and experimentation. *Gola* 22 describes the construction of a laboratory model, in the form of a *yantravalaya* (spherical device), of the rotating earth and there are allusions in the same chapter to the observer and his location. The *Pañcasiddhāntikā* has a whole chapter on astronomical instruments and methods of treating data.

6.2) a step size of $\pi/48$ allows all sines in the table to be calculated from just two principles, complementarity and the half-angle formula. But, as we shall see in a moment, this is actually of no consequence for Aryabhata's recursive construction of the higher sine differences from the two lowest sines. More directly relevant is the fact that ϵ is also the largest angle of the form π/n for which the sine is equal to the angle correct to the minute; the recursive computation could then be initiated from the inputs $s_0 = \sin 0 = 0$ and $s_1 = \sin\epsilon = \epsilon$. The arc of $\pi/48$ seems to have held a curious and mystical fascination for the Indian mathematical imagination which, also as we shall see later, defies rational explanation and which may have discouraged an appreciation of the true import of one of Aryabhata's insights and held back progress for centuries.

As for the range $(0, \pi/2)$ of the table, it just reflects the symmetry of the sine function about the angle $\pi/2$: $\sin(\pi - \theta) = \sin\theta$ (for $0 \le \theta \le \pi/2$), a relation that is blindingly self-evident when looked at 'trigonometrically', in terms of half-chords. Historically it is interesting only in that Ptolemy, if he had been aware of the symmetry, could also have stopped his table at $\pi/2$ and thus reduced his labour in half: since $\text{chord}(\theta) = 2\sin\theta/2$, the symmetry has, of course, an expression in terms of chords which is easily seen to be $\text{chord}(\pi-\theta)^2 = 4 - \text{chord}(\theta)^2$ (in the unit circle). Geometrically, this is a simple consequence of the orthogonality of $\text{chord}(\theta)$ and $\text{chord}(\pi - \theta)$ when their arcs are contiguous. However implausible it may appear that Ptolemy missed this point, it would have been even harder to miss if he had worked with half-chords (Aryabhata's half-chords, i.e., sines) instead of full chords.

The truly creative aspect of the sine table concerns point i), the conspicuous role assigned to the differences rather than the sines themselves. For an elucidation of what it tells us, an analysis of *Gaṇita* 12, the rule for the construction of the table ('Aryabhata's rule' from now on), is indispensable. A detailed look at the mathematics of the rule, including its proof, will be part of the next section. While numerous commentators, traditional and modern, are agreed on the reading of the mathematics the rule conveys, there were also serious misunderstandings and disagreements about the full significance of Aryabhata's ideas (before they were vindicated by Madhava's work). Fortunately, these ideological confusions do not affect the purely mathematical interpretation of *Gaṇita* 12. Both the mathematics and the underlying philosophy will be a sort of running theme through most of the rest of this book; for the purpose of the present section however, concerned more with the historical context and less with consequences, it is sufficient just to state the rule.

The wording of the rule is, linguistically, not unambiguous, mainly on account of the word *jyārdha* doing duty for both the sine and the sine difference. As it happens, with either meaning, the mathematics is unchanged because the formulae can be expressed either in terms of the sine differences or, by writing out the differences, the sines themselves. The commentaries came to terms with the verse each in its own manner (for the details, see the annotaed translation of Shukla and Sarma ([AB-S]). The formulation in terms of differences, preferred by the Nila school (it is for a good reason that some modern scholars call it the

Aryabhata school), is captured in the equations

$$\delta s_1 = s_1 = \epsilon, \qquad \delta s_2 = s_1 - \frac{s_1}{\epsilon} = s_1 - 1, \qquad \delta s_m = \delta s_{m-1} - \frac{1}{\epsilon}\sum_{i=1}^{m-1} \delta s_i$$

for $m = 3, \cdots, 24$.

In the historical context, these are astonishing formulae: they present a *linearly* expressed recursive solution to an essentially nonlinear problem, that of the relationship between an arc and its half-chord. As they stand, two qualifications are in order: the equations are approximate in a mild sense and they depend on the sines being measured in minutes. But, as will soon be seen when we take up the issue of where Aryabhata's rule came from, the linearity does not depend on the approximation. To repeat, what working with sine differences has achieved is to linearise, in a recursive algorithm, a fundamentally nonlinear problem. It will also become clear in due course that the particularly simple linear form of the solution is made possible by the very special nature of the trigonometric functions sine and cosine arising from their connection to the geometry of the circle. There is probably no need to add that the solution also reduces the computation to a few totally trivial steps – no square roots – but compared to the signficance of the conceptual breakthrough the saving in computational labour is as nothing.

Almost as remarkable as the rule is the fact that Playfair drew the correct conclusions about its trigonometric foundation just from the scraps of the *Sūryasiddhānta* that he had and without benefit of prior commentaries. After explaining the rule in his own geometric language, he goes on to say:

> The geometrical theorem, which is thus shown to be the foundation of the trigonometry of Hindostan, may also be more generally enunciated. "If there be three arches in arithmetical progression, the sine of the middle arch is to the sum of the sines of the two extreme arches, as the sine of the difference of the arches to the sine of twice that difference".

In equations,

$$\frac{\sin\theta}{\sin(\theta+\phi)+\sin(\theta-\phi)} \left(= \frac{1}{2\cos\phi}\right) = \frac{\sin\phi}{\sin 2\phi}.$$

This is of course a consequence of the addition theorem for sines. To go from the sum of sines to their difference, express the sum in terms of the difference of consecutive differences:

$$\sin(\theta+\phi)+\sin(\theta-\phi) = (\sin(\theta+\phi)-\sin\theta) - (\sin\theta - \sin(\theta-\phi)) + 2\sin\theta,$$

resulting in

$$(\sin(\theta+\phi)-\sin\theta) - (\sin\theta-\sin(\theta-\phi)) = 2(\cos\phi - 1)\sin\theta = -2\sin^2(\phi/2)\sin\theta.$$

This second difference formula is the key input in the derivation of Aryabhata's rule; all that is required is to specialise the variables θ and ϕ to appropriate multiples of ϵ.

Historically, Playfair was slightly off in respect of one detail. Aryabhata almost certainly did not have the addition theorem for sines. Though it follows from one of Brahmagupta's theorems on cyclic quadrilaterals (which has Ptolemy's theorem as a corollary, Chapters 8.4 and 10.2, to come), it is absent in the pre-Nila texts (with one possible and controversial exception in the form of a supposed supplement to one of Bhaskara II's books, see Chapter 10.2) and Madhava received much acclaim for discovering it. The addition theorem is actually unnecessary for Aryabhata as the equation for the second difference is a simple consequence of the formula for the symmetric difference:

$$\sin(\theta + \phi) - \sin(\theta - \phi) = 2\cos\theta\sin\phi$$

and its cosine counterpart:

$$\cos(\theta - \phi) - \cos(\theta + \phi) = 2\sin\theta\sin\phi$$

and these are very easy to prove from elementary geometry without using the addition theorem, see the next section.

Aside from acknowledging the pioneering and almost forgotten contribution of John Playfair (more diligently cited than read, it would seem) to our understanding of the revolution that Aryabhata's trigonometry brought about, there is another reason for my excursion into historiography, that of restoring a degree of balance to some of the more recent writings on the subject. Playfair's rigorous examination of Aryabhata's rule and his equally rigorous search for its mathematical justification – instead of just declaring that Aryabhata must have got his ideas from Ptolemy – are in some contrast to some of the later European and, more generally, Western work in the subject. There is in fact still an active line of scholarship, initiated by some famous names and carried forward by specialists in Babylonian, Greek and Islamic astronomy, which seeks (and, often, sees) the sources of Indian trigonometry in Greece and Alexandria, not based on commonalities of ideas and methods (in the absence of concrete facts), but on analogies and what one may call mathematical snippets.[7] The prime exhibit in support, unsurprisingly, is the chord table of Ptolemy together perhaps with the reconstructed and conjectural table of Hipparchus as an even earlier candidate. But beyond having a shared motivation in astronomy, the two tables have virtually nothing in common, either conceptually or computationally. There is no trace in Greek geometry of Aryabhata's crucial step, the shifting of the focus from the full chord to the half-chord. Similar is the case with the other key idea, the introduction of differences and second differences as a means of linearising the arc-chord relationship. I will argue presently that

[7]For an idea of how it is done and for references, see the relevant chapters of a recent book: Glen van Brummelen, *The Mathematics of the Heavens and the Earth: the Early History of Trigonometry*, Princeton University Press, Princeton and Oxford (2009).

the roots of that particular idea are to be found in the Indian attachment to the rule of three – essentially quantitative and arithmetical – and in the recursive treatment of deviations from it as a principle of mathematical reasoning.

Independently of the sine table, Hipparchus has also been invoked as the source of Aryabhata's standard radius $R_0 = 3438$ minutes (see van Brummelen, cited above, for the history of these speculations). The credibility of the suggestion should depend ideally on reasonable evidence of transmission (which is currently lacking) but, in this instance, we have also to ask what Hipparchus could have had to transmit. Aryabhata's $R_0 = 21,600/2\pi$ comes out of measuring the radius in minutes and even then it is not a number that is immutably fixed; its numerical value at any historical time (and place) is contingent on the value of π as known at that time and place and it is a surprise that it should at all be put forward as an indicator of transmission (across a time-gap of seven centuries). For Hipparchus to have hit upon the number 3438, he would first have had to think of measuring the radius in units of arc-minutes and there is not even hearsay evidence for that; if he did, why would Ptolemy, who knew his work well, not have adopted it instead of his own awkward linear units for chords? Hipparchus would then have needed a precise enough number for π, with an error less than one part in 3438. Not knowing what value he did have, we can try Archimedes's rational bounds $223/71 < \pi < 22/7$ from a century earlier and conclude only that the corresponding R_0 should have been between 3425 and 3445. Ptolemy's $\pi = 377/120$ would have given almost the right number but we definitely know that he did not measure the diameter in angular units. The point is that 3438 is not a sacred number, it is an artefact of conventions, a misleading mathematical tidbit; the sacred number is π.

7.4 The Making of the Sine Table: Aryabhata's Rule

It is time to get back to the trigonometric origins of Aryabhata's rule. As with the other enigmatic *sūtra*s of the *Āryabhaṭīya*, our best guide to his thinking is his last great commentator Nilakantha and his (almost last) masterpiece, the *Āryabhaṭīyabhāṣya*. In style and presentation it is expansive and in lucid prose, the exact opposite of the *Āryabhaṭīya*. The mathematical parts are explained with great clarity and pedagogical skill – Nilakantha was after all a wise old man of about eighty when it was composed.

As far as the sine table is concerned, Nilakantha's method of presentation is to first give the geometric proof of the exact formula for the second difference of sines for the canonical 24-fold division of the quadrant and then to get to Aryabhata's rule by making two minor adjustments. It is true that Nilakantha was writing with the benefit of having lived after Madhava. Madhava's infinitesimalisation of the method of differences had brought about such a great deepening of Aryabhata's first insights that it would be unnatural if it did not influence the way succeeding generations understood Aryabhata's ideas. It is then a legitimate worry that Nilakantha's proof of the difference

formula may not be a faithful retracing of Aryabhata's own path to it. There is circumstantial evidence, however, that in all essential respects – we cannot of course be sure of every detail – it is. Firstly, the mathematics has nothing that is infinitesimal; the geometry is completely elementary and the proof requires no more than the time-tested technique of comparing naturally constructed similar right triangles. (To put things in perspective, it is much less subtle and demanding than the proofs of Brahmagupta's theorems on cyclic quadrilaterals). It is also a well-attested fact, from the themes and the manner of their treatment as well as from long-held historical tradition, that it is on the banks of the Nila that Aryabhatan astronomy and mathematics finally found their true home. The fidelity and reliability with which oral teaching and learning transmitted knowledge over many generations is an additional guarantee that what was attributed to Aryabhata was probably Aryabhata's. In a sense the answers to such doubts do not matter much. If a certain mathematical result is correct and is not self-evident even to the intelligent – as Aryabhata's rule certainly was not – a reasonable default position (always keeping in mind that one person's logic may be another's intuition) is that it was arrived at by a process of logical reasoning; and, as we have seen, Aryabhata would have bristled at the idea of divine inspiration.

The account below of the geometry of sine (and cosine) differences deviates slightly from Nilakantha's in one respect. His explanation is tailored from the start to the step size of $\epsilon = \pi/48$ as in *Gaṇita* 12; all angles are multiples of ϵ and, reverting to our notation $s_m = \sin m\epsilon$, the idea is to evaluate geometrically the difference $\delta s_m = s_m - s_{m-1}$ as the symmetric difference about $(m - 1/2)\epsilon$. The proof holds without change for the difference of general angles; Aryabhata very likely knew this and dealing with the general formula has the practical advantage that it comes in handy for future use.

In the diagram of Figure 7.4, the arc XY defines a quadrant of the unit circle with centre O and with all angles measured from X. A, B and C are points on it corresponding to angles θ, $\theta - \phi$ and $\theta + \phi$ respectively. The half-

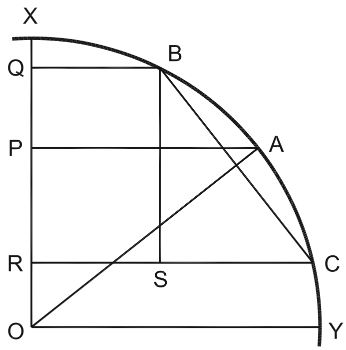

Figure 7.4: The geometry of sine and cosine differences

chords AP, BQ and CR are their sines and OP, OQ and OR their cosines. BS is drawn perpendicular to CR. Since $BQRS$ is a rectangle, we have then $CS = \sin(\theta+\phi) - \sin(\theta-\phi)$ and $BS = \cos(\theta-\phi) - \cos(\theta+\phi)$. Now, the triangles BCS and AOP are similar since the pairs of lines (BC, AO), (CS, OP) and (BS, AP) are mutually orthogonal. Hence,

$$\frac{BS}{BC} = \frac{AP}{OA}, \qquad \frac{CS}{BC} = \frac{OP}{OA}.$$

But $OA = 1$, $AP = \sin\theta$, $OP = \cos\theta$ and $BC = 2\sin\phi$ and we have immediately the familiar formulae for the symmetric differences of sine and cosine which I write again:

$$\sin(\theta+\phi) - \sin(\theta-\phi) = 2\sin\phi\cos\theta,$$

$$\cos(\theta+\phi) - \cos(\theta-\phi) = -2\sin\phi\sin\theta.$$

Specialise now to $\theta = (n - 1/2)\epsilon$ and $\phi = (1/2)\epsilon$ to get the canonical differences

$$\delta s_m = s_m - s_{m-1} = 2\sin\frac{\epsilon}{2}\cos\left(m - \frac{1}{2}\right)\epsilon,$$

and

$$\delta c_m = c_m - c_{m-1} = -2\sin\frac{\epsilon}{2}\sin\left(m - \frac{1}{2}\right)\epsilon,$$

where $c_m = \cos m\epsilon$.

Faced with this pair of coupled linear inhomogeneous equations, the modern reader will know what to do: substitute one equation into the other. In other words, take the second difference of the sines

$$\delta s_m - \delta s_{m-1} = 2\sin\frac{\epsilon}{2}\left(\cos\left(m - \frac{1}{2}\right)\epsilon - \cos\left(m - \frac{3}{2}\right)\epsilon\right)$$

and use the cosine difference formula. The result is

$$\delta s_m - \delta s_{m-1} = -4\sin^2\frac{\epsilon}{2}s_{m-1}.$$

That is what Nilakantha (and probably Aryabhata) did except in one respect. Instead of breaking up the geometric reasoning separately for sine and cosine and then bringing them together algebraically, he does the substitution geometrically so to say by deriving the cosine difference from the beginning in parallel with the sine difference, not for the right triangle with vertices at $(m + 1/2)\epsilon$ and $(m - 1/2)\epsilon$ as I have done here, but for the one with vertices at $(m - 1/2)\epsilon$ and $(m - 3/2)\epsilon$ (Nilakantha's geometric substitution is described in [AB-S]). It is also easy to check, though unnecessary for the sine table, that the cosines satisfy the identical equation

$$\delta c_m - \delta c_{m-1} = -4\sin^2\frac{\epsilon}{2}c_{m-1}.$$

Nilakantha now uses the lowest second difference $\delta s_2 - \delta s_1 = -4\sin^2(\epsilon/2)s_1$ to eliminate the half-angle factor on the right:

$$\delta s_m - \delta s_{m-1} = \frac{s_{m-1}}{s_1}(\delta s_2 - \delta s_1)$$

for $m > 1$ (for $m = 2$ it is empty). This equation is exact and it is valid for any step size of the form $\pi/2n$.

There are several ways in which the equation can be reexpressed. The one Aryabhata chooses relies on the fact that the sum of the successive differences of any ordered sequence, not necessarily equally spaced, is the difference between the two extreme elements: $\delta s_{m-1} + \delta s_{m-2} + \cdots + \delta s_1 = s_{m-1} - s_0 = s_{m-1}$ in our particular case, so that the exact sine-difference equation can be rewritten as

$$\delta s_m = \delta s_{m-1} - \frac{(\delta s_1 - \delta s_2)}{s_1} \sum_{i=1}^{m-1} \delta s_i.$$

Once the values of s_1 and $\delta s_1 - \delta s_2 = 2s_1 - s_2$ are known, the equation determines each δs_m recursively from the lower differences and, thereby, each s_m recursively from the lower sines. The recursive structure is independent of these 'initial values' as is, equally obviously, the linear structure, the latter being a consequence of the fact that the first differences of sines and cosines are proprtional respectively to the cosines and sines at the mean points. A final remark is that this as well as the apparently trivial but very general property of differences, that the parts so defined sum up to the whole, are the seeds which Madhava and the Nila school nurtured into the infinitesimal calculus of trigonometric functions.

How did Aryabhata go from the exact formula to the rule as given in the previous section? The rule itself, confirmed by the numbers in *Gītikā* 12, tells us that $\sin\epsilon$ was taken to be ϵ. The only question is about how it was verified to be a good approximation and the unanimously accepted answer is that it was computed by applying the half-angle formula to go from $s_8 = 1/2$ to s_4 to s_2 to s_1 which last is 225 minutes, equal to ϵ to within 1 minute. Since one has to compute s_2 along the way, the other constant $\delta s_1 - \delta s_2 = 2s_1 - s_2$ is also thereby determined and it turns out to be 1 minute and is therefore absent in the numerator of the second term. This accidental simplification due to the use of minutes to measure chords can be and has been a source of occasional confusion.

It is clear at the end of it all that the only place the rule deviates from exactness, in fact the only place it is sensitive to the choice of the step size, is in the constants specifying the initial conditions, the two adjustments referred to earlier.

The position Aryabhata's table came to acquire as an integral part of Indian astronomy has few parallels in any branch of learning. Virtually every subsequent work, beginning with Varahamihira and until Shankara Varman (the last descendent of the Nila lineage, early 19th century, different from the 16th

century Shankara Variyar) had a sine table, not always of sine differences, and of varying precision. In many texts from Kerala, beginning in the 9th century (well before the Nila period), the same table appears in identical *kaṭapayādi* notation, as though it was formulated once and for all, to be passed on to the next generation like some sort of secular *mantra*. Elsewhere and at other times, the method of computation and sometimes even the step size varied. Among the earliest, Brahmagupta got away with sines in steps of 15 degrees, very easily computed, and that is because he had a powerful interpolation formula for filling in the gaps. Varahamihira (see Chapter 6.2) and Bhaskara I also preferred the theoretically far less interesting half-angle formula. The *Sūryasiddhānta* apparently could not make up its mind: it describes Aryabhata's rule in terms of differences but its table is for sines, possibly because that is what is needed in astronomy. Outside astronomy, Bhaskara II put the table to use as input data in a computation of the correct surface area and volume of the sphere by a process of finite numerical integration.

But the more important reason for my going over the genesis of the table in such, perhaps pedantic, detail is to lay out its different mathematical components, to redraw attention to the very original strands of thought that were woven into the fabric of Aryabhata's trigonometry and, especially, to highlight the unexpected new directions in which the introduction of the trigonometric functions took the geometry of the circle. This was a historical landmark and brought with it the recognition that the sine and the cosine are proportional to each other's finite differentials. To these we can add yet another first, not in principle restricted to geometry: the use of discrete (finite) differentiation followed by discrete integration as a method of treating nonlinear problems. Aryabhata probably was aware of the generality and depth of these novel ideas and methods – he was not an unduly modest man, as the flaunting of "the boat of my own intelligence" shows. But we cannot say the same about those who followed him. The subtlety of his mathematical ideas caused much confusion and misunderstanding among them – in contrast to the astronomy, with the obvious exception of the spinning earth – and for a long time, until Madhava arrived on the scene. That story is interesting enough to be taken up for a deeper examination in the next (and concluding) section of this chapter. It will also give us the opportunity for a final summing up of what constitutes or should constitute the discipline of trigonometry and how it is distinct from the geometry of the circle as it was before Aryabhata.

7.5 Aryabhata's Legacy

Gaṇita 12 is the last in a sequence of four interlinked verses. The second line of the first of them, *Gaṇita* 9, states that the chord of $\pi/3$ is the radius, i.e., $\sin\pi/6 = 1/2$, the starting point for computing the initial values s_1 and $\delta s_2 - \delta s_1$. The next verse is famous for its value of $\pi = 3.1416$, qualified as *āsanna*, which does not quite mean 'approximate' as it is often translated

(there are other words for it, e.g., *sthūla*) but 'proximate', close but not exactly equal to. It is this word that Nilakantha takes as the cue for his assertion that π is an irrational number or, in the Greek manner, that it represents an incommensurable ratio: there is no unit that measures both the circumference and the diameter without leaving a remainder, however small, in at least one of them. (There is a somewhat vague attempt to come to terms with *āsanna* by Bhaskara I, comparing the "inexpressible" π with $\sqrt{10}$). It is worthy of note that the very assertion is premised on the acceptance of a common but arbitrarily chosen unit, not restricted to $\pi/48$ or even a minute of arc, to measure arcs and chords in. The other message of *Gaṇita* 10 is that a sine table cannot be made without a value for π, that an accurate sine requires a correspondingly precise π *via* its effect on the standard radius R_0. The message did not always get across to the later table-makers.

In contrast to the clarity of these verses is the intriguing *Gaṇita* 11. A convincing reading of it will, we can hope, open a window to Aryabhata's thinking on how to divide the circle into arc segments and how then to approach the problem of computing their sines. Linguistically, the nouns occurring in the verse – quadrant, half-chord, trilaterals and quadrilaterals, etc. – and the one verb – to cut up or divide – are clear enough. Problems arise with the qualifiers, specifically one word, *yatheṣṭāni* ('according to one's wish' or 'as one pleases') and what it applies to. Commentators from very early on have been unsure on this question, leading to dramatically different interpretations of what particular mathematical idea was intended to be conveyed.

The most natural reading grammatically is that it is an adverb modifying the verb divide – *chindyāt . . . yatheṣṭāni* – and with that interpretation the stanza has a fluent but still faithful translation:

> Divide equally, according to one's wish, a quadrant of the circle. The half-chords of equal arcs [so formed] [are determined] from the radius by [means of] trilaterals and quadrilaterals.

What is its mathematical reading? Bhaskara I takes it as referring to the half-angle construction of the table, and many commentators old and new have followed in his footsteps. Looking at it literally, the connection is almost invisible; one searches in vain for a link between any step in the half-angle geometry and any phrase in the stanza (except of course for the inescapable arcs and chords and trilaterals). Specifically, there are no quadrilaterals in the half-angle construction (see Figure 6.1). Still more seriously, it does not work for arbitrary ('as one pleases') divisions of the quadrant, only for divisions of $\pi/6$ and $\pi/2$ by powers of 2.

There is another possibility, somewhat far-fetched: Bhaskara I's approach has nothing to do with *Gaṇita* 11 and 12; its source is different, perhaps Aryabhata's supposedly lost book. (Mystery: why is it then part of Bhaskara's explication of *Gaṇita* 11?). That will also explain the simultaneous popularity of both the sine table (the half-angle method is not mentioned in *Āryabhaṭīya*)

and the sine difference table in the early commentaries, in one case in the same book (*Sūryasiddhānta*).

Gaṇita 11 makes excellent sense however if we take it as preparing the ground for Aryabhata's rule for sine differences as presented in the next stanza. Every element in the verse fits in with such an interpretation if *jyārdha* is taken as sine difference rather than sine but that, in any case, is its meaning in the operative stanza *Gaṇita* 12 and the numerical table of *Gītikā* 12 (and elsewhere). And quadrilaterals (rectangles) arise naturally in the geometry of the sine difference formula, in fact are essential ($BQRS$ in Figure 7.4). But, beyond that utilitarian interpretation, the use of *yatheṣṭāni* gives the verse an altogether new dimension: the derivation of sine and cosine differences as the basis of Aryabhata's rule is valid for arbitrary, in particular arbitrarily small, angular steps, with only minor adjustments to be made in the initial values. (The approximation $\sin\epsilon = \epsilon$ only gets better as the step size is decreased). The authors of the Nila school certainly knew this as it is the first step in the development of the power series for the sine. My reading of *Gaṇita* 11 would suggest that Aryabhata already knew it and meant it; it is simple enough, the geometry does not need any modification. Seen in this light, the verse is the opening announcement of a programme of which *Gaṇita* 12 represents the first stage, a vision statement.

Independent confirmation that Aryabhata had in mind the notion of an infinitesimal division of the circle comes from the form in which he expressed the area of the circle as the product of half the circumference and half the diameter (in *Gaṇita* 7). Turning once again to Nilakantha, we find its justification in the area equivalence of the circle with a rectangle of sides $C/2$ and R, established by a cut-and-paste method applied to the regular inscribed $2n$-gon, in the limit of large n. The radii to the vertices divide the polygon into $2n$ congruent isosceles triangles of base $P/2n$ and side R (P is the perimeter of the polygon). Cut out and paste two of them with the radial sides aligned and the vertices opposed so as to form a parallelogram (Figure 7.5). As n is made larger and larger, the polygon tends to the circle, P tends to its circumference C, each infinitesimal parallelogram to a rectangle of sides R and $C/2n$ and hence the total area to $nRC/2n$. To be noted is that there is no finitistic cut-and-paste that will do the job because that will amount to squaring the circle, given the *Śulbasūtra* constructions of area-equivalent squares and rectangles. The same

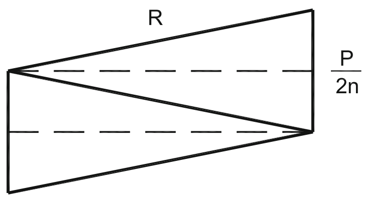

Figure 7.5: Infinitesimal elements of the area of the circle

argument with a slight change is found also in *Yuktibhāṣā* which then uses very similar cut-and-paste reasoning as the basis of its calculus method, with the limit carefully handled, for the area and volume of the sphere.

The idea that trigonometry can be done with arcs as small as we please and that an arc is not the same as its chord, no matter how small, had an amazingly uncomprehending and hostile reception already from Aryabhata's first apostle, Bhaskara I. Only 130 years after the *Āryabhaṭīya*, he wrote in his *bhāṣya* on *Gaṇita* 11 (the translation is that of Plofker[8]):

> It is proper to say that a unit arc can be equal to its chord; even someone ignorant of treatises knows this; that a unit arc can be equal to its chord has been criticized by precisely this [master].
>
> But we say: An arc equal to a ([its]) chord exists. If an arc could not be equal to a ([its]) chord then there would never be steadiness at all for an iron ball on level ground. Therefore we infer that there is some spot by means of which that iron ball rests on level ground. And that spot is the ninetysixth part of the circumference.

From a historiographic and epistemic point of view, the passage is invaluable, never mind the wrong mathematics (and physics; Bhaskara I could have drawn a big enough circle on the ground and checked for himself). It is confirmation, from the horse's mouth as it were and in strong words, that what we have so far inferred from internal evidence was in fact the master's own view; he was in fact speaking of arcs smaller than $2\pi/96$. That it occurs specifically in the commentary on verse 11 further strengthens the conclusion that *yatheṣṭāni* has exactly the mathematical sense we have read into it, 'arbitrarily small'. The sharpness of the disagreement with the venerated master is perhaps a measure of how deeply embedded was the idea of a circle as a many-sided regular polygon and of the inability to envisage it as the limit of the polygon as the number of its sides grew without bound. This is the same issue that came up in connection with Brahmagupta's (and Bhaskara II's) struggle with division by zero and the reciprocal relationship of zero and infinity (Chapter 5.4). As was noted there, the progression from discrete to continuous (if not continuous, at least 'sufficiently small') change, from the finite to the infinitesimal, has been a struggle in the Greek-European tradition as well, tied up intimately with the story of calculus. In India and in the context of the calculus of the circle, the struggle was finally resolved by Madhava in a very pragmatic but rigorous manner, without having to introduce the general notion of what we now call real numbers: divide a geometrical quantity by a large number N and let N increase without bound – exactly as in Aryabhata's area of the circle – making sure all the while that simplifications made along the way become exactly valid in the limit.

But before that happened, the idea that $2\pi/96$ is a sort of quantum of angle, the unit below which one may not venture, enjoyed a surprisingly long

[8] Kim Plofker, "Mathematics in India" in [MEMCII].

life. We do not know where the idea of its indivisibility came from but at some point it acquired religious sanction: a text named *Brahmasiddhānta* (not one of Varahamihira's Siddhantas, possibly part of the Puranic corpus, from around or soon after Aryabhata's time) says: " The ninetysixth part of the circle is straight as a rod". In the centuries of the return to Hindu orthodoxy that followed the terminal weakening of Buddhism, notions having religious or quasi-religious backing tended to live long. Anyhow, no one seems to have calculated sines of angles smaller than $2\pi/96$, though approximate formulae for interpolating between the canonical (or larger) steps were derived, by Brahmagupta among others. The contradiction is difficult to explain rationally.

A judgement on Aryabhata's impact on Bhaskara I's contemporary Brahmagupta has to be somewhat more nuanced. Brahmagupta's doctrinal disapproval of Aryabhata, expressed freely on more than fifty occasions in the *Brāhmasphuṭasiddhānta*, need not in principle have come in the way of a proper appreciation of the mathematics itself. In fact as well, he did not repudiate any of Aryabhata's mathematics, not even the wrong volume formulae; trigonometry plays its full part in his astronomy which, as far as the technical aspects are concerned, differs from Aryabhata's only in detail. Nevertheless, his groundbreaking theorems on cyclic quadrilaterals owe little to Aryabhatan geometry. More to the point, he uses the old value of $\sqrt{10}$ for π and the interpolation formula for the sine is presented (in his other book, *Khaṇḍakhādyaka*, about 30 years after the *Brāhmasphuṭasiddhānta*) in the context of his own table, with a step size of 900 minutes; the significance of *yatheṣṭāni* seems to have been lost on him or, perhaps, he felt that with an accurate enough (2nd order) method of interpolation, there was no need to insist on a very fine division of the quadrant to start with. On the other hand, his treatment of *kuṭṭaka* is Aryabhatan in all essentials and is introduced by invoking the question of determining the *yuga* "for two, three, four or more planets from their elapsed cycles": the same problem and the same solution. And the problem of the quadratic Diophantine equation, whose study he pioneered, is clearly a mathematical generalisation from Aryabhata's linear problem (it has no direct astronomical use). Perhaps we should take a sentence at the beginning of *Khaṇḍakhādyaka*, that its results are identical with those of "*ācārya* Aryabhata", as the homage of an older and wiser Brahmagupta to his already illustrious predecessor.

As in the case of Brahmagupta, a balanced assessment of the extent to which Bhaskara II was influenced by Aryabhata is best allowed to emerge after we go over the high points of his achievements in some of the chapters to come. The following paragraphs may therefore be taken as a general summing up, to be read again later if the reader wishes. About Bhaskara's commitment to Aryabhatan ideas there has never been any doubt. Almost all of his astronomy and much of the mathematics are developments of material first found in the *Āryabhaṭīya* and directly inspired by it. Still, there are puzzles which may make one wonder whether all of Aryabhata's ideas, especially regarding trigonometry in the small, found a receptive mind in Bhaskara II. The seeds of doubt sprout most visibly in his treatment of the surface area and volume of the sphere,

although he gave their correct formulae for the first time in India. His method can be accurately described as discrete integration and the point of interest here is that his division of a great circle, which is the first step in the computation, is resolutely the canonical 96-fold one. The total area (and the volume similarly) is decomposed into area elements which are treated as though they are planar figures, ignoring the effects of curvature. The computation is then reduced to adding up the numerical areas of these (finite number of) elements, each of which is proportional to $(1/2)(s_m + s_{m+1})$.

Despite the promise of the initial step, that of a decomposition into small elements, the follow-up is disappointingly routine. The sum of the canonical sines $\sum_1^{24} s_i$ (apart from an easy end correction) is performed by looking up the numerical value of each s_i from the table and just adding them up; his namesake from five centuries earlier might have proceeded in exactly the same manner. The essential point of Aryabhata's listing of differences is that each sine is the finite differential of the corresponding cosine (and conversely) or, equivalently, the second differential of itself; Bhaskara could have reduced his labour effectively to zero by using the main insight of Aryabhata's finite calculus, that $\sum_1^{24} s_i$ is proportional to $c_{24} - c_0$ (again, ignoring some notational subtleties, see below). That would also have improved the numerical accuracy but since he knew the exact formula, it probably did not matter to him.

That brings up a second puzzle: how could such an approximate numerical procedure have led to the exact algebraic formula for the area of the spherical surface? It is easy enough to add up the tabulated sines and find that the sum is numerically not the same as $21600 \times 2 \times 3438$ of the formula, by a significant enough margin. One possible answer to the question is that he arrived at the formula by a method other than numerical and the calculation was meant to be no more than a verification, a numerical experiment to check a theory. If that was indeed the case, he has left no word of what that theory might have been. What we read in his own expalanatory notes (known as *Vāsanā*) to the *Siddhāntaśiromaṇi* is a careful description of the method of decomposition of the surface followed abruptly by the formula with no further explanation (the summing of the sines comes later); that is unlike Bhaskara who loved to expound, and not very enlightening for us.

There is a plausible resolution of these puzzles. That is to suppose that Bhaskara II knew perfectly well that a step of $2\pi/96$, or any finite step for that matter, will not lead to the exact formula, that better and better approximations will result from progressively reducing the step size, but to suppose also that he did not have the technical means to carry it through to his satisfaction. At one point in the explanation of how to divide the area of the sphere by latitudes, the *Vāsanā* gloss says in passing that more than 24 half-chords can be used to create a finer division. That would have required the making of a finer sine table if the sines are just summed. If on the other hand the sine is replaced by the cosine difference, one can use the exact formula $\sum_1^n s_i = c_{n+1/2} - c_{1/2}$ (rather than the schematic $c_n - c_0$ of my earlier paragraph) which is true for

an arbitrary step size $\pi/2n$. If finally the cosines at the end points are rounded off as $c_{n+1/2} = c_n = \cos\pi/2 = 0$ and $c_{1/2} = c_0 = 1$, what results is the correct algebraic formula for the area, independent of the step size. The suggestion then is that this was Bhaskara's theory for areas and volumes of spheres, and that the choice $n = 24$ was out of respect for convention. The justification of the rounding off, valid strictly only in the limit $n \to \infty$, would have been the sticking point. The passage to the limit, as we know, was not understood and implemented until Madhava arrived on the scene.

By the time of Bhaskara II, the 33 verses of the *Gaṇitapāda* had become the driving force behind much of the mathematics done in India, with the partial exception of Brahmagupta's cyclic geometry. Apart from those like the two Bhaskaras who wrestled with the more innovative, infinitesimal elements of it (Brahmagupta kept away from them), there were others who wrote books, taught students and otherwise kept the flame alive. To mention only a few: Śrīdhara the author of *Pāṭīgaṇita* ('Writing-board Mathematics'); Śrīpati who worked on all branches of mathematics and astronomy; the learned Pṛthūdaka who, while writing the definitive commentary on Brahmagupta's work, vigorously defended Aryabhata against the criticisms of his own intellectual mentor; Mahāvīra from south Karnataka, a Jaina and the last representative of the old Jaina tradition, possibly the least Aryabhatan, trigonometrically speaking, among them; and Aryabhata II who, like Mahavira, made significant contributions to the algebra of pulverisation but also had difficulties with the geometry. To these have to be added the names of Govindasvami and his disciple Shankaranarayana, devoted followers of the Aryabhatan doctrine *via* the teachings of Bhaskara I, as the first astronomers outside legend and myth to have lived and worked in Kerala. It seems fair to think of them as teachers and consolidators rather than as creative mathematicians of the first order.

Aryabhata's work has the distinction of being the first historically authenticated instance of Indian mathematics and astronomy travelling beyond India's cultural boundaries. It came about not directly – it is intriguing to imagine what anyone not within his direct sphere of influence would have made of the *Āryabhaṭīya* – but through the writings of Brahmagupta. Brahmagupta's life span coincided with the rise and rapid spread of Islam in western Asia (and, very soon after, beyond) and the reawakening of the long-dormant mathematical inheritance of Mesopotamia through the interest and support of the Caliphate in Baghdad. But what the Islamic scholars discovered, at first, was not their own legacy but the sciences of Hind. Brahmagupta's two books were among the first to be studied and translated into Arabic, followed later by the *Āryabhaṭīya*. The story is well known; there are very few (if any) other examples as well documented of the dissemination of a whole body of knowledge across cultural borders in the period called medieval in European history. The contrast with conjectural models of putative earlier transmissions that we have met before, based at best on loose analogies, cannot be sharper.

Less widely known is that, during roughly the same period, India-China contacts – religious, cultural and commercial – of earlier centuries had bur-

geoned and led to a flow of Indian astronomical knowledge and of its attendant mathematics eastwards as well. We have already noted that Aryabhata's π may well have had a Chinese gene in it. In the 7th century and later a significant number of Indian men of science, generally but not always with a Buddhist background, travelled to China and lived there for long periods; three of them are recorded as having worked in the official Bureau of Astronomy, one as the Director. Aryabhata became known in learned circles in China and the sine got a Chinese name. The Chinese opening to India came to an end in the 11th century with the decline in interest in Buddhism. That was about the time when Islamic-Indian mathematical astronomy, having earlier arrived in Spain, began to impact Europe. Another irony of history.

What do we know about the sources from which Aryabhata himself might have drawn inspiration? We have already gone over the evidence for the direct role of Hellenistic Alexandria in his astronomical model and, at a more abstract level, in the idea of astronomy as a mathematical and predictive science. There is plenty of it, textual and circumstantial; there is even praise in Varahamihira's astrological book, the *Bṛhatsaṃhitā*, for the *yavana* astrologers/astronomers even as he looks down on them because they were *mleccha*, beyond the pale. There is however no external documentary support at all for an accompanying flow of purely mathematical ideas into India. As for internal evidence from the mathematical substance of the *Gaṇitapāda*, it is not unreasonable to attribute the Euclidean algorithm of the *kuṭṭaka* to Greek influence (though the Greeks do not seem to have had a similar use for it; no *yuga* in Alexandria) on the basis of priority as well as its absence in India before the *yavana* Siddhantas. But beyond that, particularly where geometry is concerned, the *Gaṇitapāda* owes nothing to Greece: the conceptual foundation is different and the results and methods show a parallelism only in the knowledge of those universally valid truths of mathematics that were known to both cultures. That, hopefully, is clear enough from the details covered in this chapter.

To anyone familiar with the *Śulbasūtra*, the indigenous geometric inspiration should be just as clear. The circle with mutually perpendicular intersecting chords and diameters was a motif already in the Indus culture (ca 2500 BCE) and has been a staple of Indian geometry from the beginning of the 1st millennium BCE. (I like to imagine that the idea of the half-chord came from the *Śulbasūtra* determination of the cardinal directions; compare Figures 2.1 and 7.3). Likewise the theorem of the diagonal in its *Śulbasūtra* form, as holding for right triangles or rectangles of arbitrary real (as we would say now) sides. The general geometric theorem, as distinct from the existence of rational Pythagorean triples, is vital for Aryabhata's synthesis. Not only does it metricise the geometry of the circle (and, eventually, all geometry); it also brings in the notion of the sine as a (real, periodic) function of the angle. In the light of *yatheṣṭāni*, that is what the sine table is about: take a value of the angle (ideally, arbitrary) as input and generate its sine as the output. If trigonometry is the analytical study of the geometry of the circle, the *Āryabhaṭīya* marks its birth.

The question of when a new mathematical discipline took form and established itself as such has always been a legitimate concern of historians, seldom settled to everyone's satisfaction. A first criterion surely is that it must, by transcending existing modes of thought, enable the elucidation of issues and the solution of problems previously inaccessible or not thought of. Generally speaking, that will entail the introduction of fresh concepts and the fashioning of new tools which, in turn, will open up new avenues of enquiry, becoming in due course a well-demarcated mathematical domain with its own distinctive characteristics. The trigonometry of the *Gaṇitapāda* qualifies on all these counts, though the European recognition of the individual identity of the new discipline and the consequent invention of a new name had to wait until the 15th or 16th century, without, as far as we know, any direct knowledge of the *Āryabhaṭīya* or its precise contents. The name trigonometry itself is a reflection of this ignorance, stressing as it does one particular aspect of the subject: being really about the circle (think of the theory of periodic functions), trigonometry can be profitably used to study the geometry of cyclic figures and all triangles are cyclic. Every proposition concerning such figures has, in principle, a trigonometric formulation and a trigonometric proof. (We shall come across a very interesting example of this mutability, from *Yuktibhāṣā*, in Chapter 10.2). From this to the rewriting of a few adaptable theorems of Classical geometry (and some of the early astronomical estimates) in terms of sines and cosines as done by some modern historians, and to the declaration that trigonometry began with Euclid if not earlier (Pythagoras?), is a temptingly easy step. Perhaps we should call all early circle and sphere geometry 'trigonometry without trigonometric functions'; after all, everything can ultimately be traced back to the fact that the locus of points equidistant from a fixed point, all in the same plane, subtends four right angles at it, independent of the radial distance.

Not too much attention is generally paid to the antecedents of the final step of Aryabhata's approach to the sine table, namely, the resort to differences and second differences; certainly, no Greek precedents have been pointed out. For an appreciation of where the idea might have come from and where later it led to, it is helpful to adopt, temporarily, a more general and modern viewpoint. Thus, consider a (real) function f of a (real) variable x on the interval [0,1] which is 'good' enough (continuous with a continuous derivative will do for the purpose here). Let $x_0(=0), x_1, \cdots, x_n(=1)$ be equally spaced points – i.e., to anticipate later developments, divide the interval by n – and denote $f(x = x_i)$ by f_i. If f is the constant function, the difference $\delta f_i = f_i - f_{i-1}$ vanishes for all i. If f is linear in x, the rule of three (extended to more than two ratios) holds in the form: $\delta f_i / \delta x_i$ (where $\delta x_i = x_i - x_{i-1}$) is the same for all i; i.e., since $\delta x_i = 1/n$, δf_i is constant. Consequently, the second differences $\delta^2 f_i = \delta f_i - \delta f_{i-1}$ vanish. (This is all very elementary; my purpose is only to establish a *paribhāṣā*, a glossary to link the way we think of such things to notions Aryabhata was familiar with). When the rule of three is violated and δf_i is not constant, the dharma of recursive refining is to start with the linear approximation as the first guess and to seek to determine the deviation from

linearity by looking at the 'error' as measured by the second difference (and so on; in general the process will not terminate).

The rule of three is almost as ancient as geometry itself in India, going back to the ideas of scaling and similarity of triangles in the *Śulbasūtra*, and the method of recursive corrections may also be as old (the *Śulbasūtra* $\sqrt{2}$). The debt Aryabhata owes to these notions and techniques is perfectly obvious and his treatment of the function sine (and cosine) in its domain $[0,\pi/2]$, with $n = 24$, can serve as a textbook illustration of their generality and interconnectedness. The miraculous property that made possible a complete determination of these functions (at chosen equally spaced angles) is also clear: their second differences are proportional to the functions themselves, making it unnecessary to go to higher differences. From the perspective of the rule of three, these functions are thus nonlinear in the simplest possible way. None of this is a surprise for the modern reader: the miraculous property is that the derivatives of the sine and the cosine are proportional to each other and their second derivatives are their own negatives. This last, as also we know, translates into geometry as the statement that the circle is the curve with constant (negative) curvature. The recognition and written acknowledgement of the relationship between the second derivative and curvature as a manifestation of the deviation from linearity, *via* the rule of three, had to wait until the implications of *yatheṣṭāni* had led Madhava to the infinitesimal calculus of the trigonometric functions arctangent, sine and cosine. But when it did come, it could not have been clearer. The very brief section on tabulated sines in *Yuktibhāṣā* says, after warning that the use of the rule of three for arcs and chords will lead to gross error:

> The reason for this: the second arc is twofold the first arc. ... The second chord is not twofold the first chord, the third chord is not threefold, and so on. The reason for this: the first arc has no curvature and is almost equal to the chord since the *śaram* (the arrow, $1 - \cos$) is small. So do not apply the rule of three to the arc because the result will be gross.

Thus was the ghost of Bhaskara I's fallacy finally exorcised.

A last remark. Functional differences have the property of adding up to the difference of the values of the functions at the end points:

$$\sum_{i=1}^{n} \delta f_i = f_n - f_0$$

or, more generally (when f is a sufficiently good function),

$$\sum_{i=1}^{n} \delta^k f_i = \delta^{k-1} f_n - \delta^{k-1} f_0$$

even if the points x_i are not equally separated. The various rearrangements and simplifications of Aryabhata's rule are based on astute applications of this

elementary principle. Trivially obvious as it is, it is the discrete form of a result which, when the limit $n \to \infty$ is taken for equal separation and with the necessary care, becomes the fundamental theorem of calculus in its basic form:

$$\int_0^1 \frac{df}{dx} dx = f_1 - f_0.$$

A historical footnote is that, in Europe much later, it is through finite functional differences that Leibniz initiated his approach to calculus for general functions and curves;[9] he could not have missed the fundamental theorem as indeed he did not. Its discrete version may be just common sense – the whole is the sum of its parts – but it is another of Aryabhata's legacies the full realisation of whose latent power lay far in the future. In India, it comes into its own, once again, in the work of the Nila school; the amount of trouble *Yuktibhāṣā* takes in explaining how the discrete fundamental theorem is employed, supplemented by a rare illustrative example, is a tribute to its importance in the context of calculus.

In the long history of the sciences in India there is only one work which surpasses the *Āryabhaṭīya* in the depth and durability of its influence and that of course is Panini's *Aṣṭādhyāyī*. The two have many striking similarities: the scope of their ambition, the universality of the abstract conceptual framework they each put in place and the economy and versatility of the means employed. When Nilakantha's magisterial commentary makes a reference to itself as *Mahābhāṣya*, like Patanjali's 'Great Commentary' on the *Aṣṭādhyāyī*, we do not have to wonder what precedent he had in mind. Thanks to the great resurgence of interest in the formal study of language that began in Europe in the second half of the 19th century, Panini's achievements are now universally acknowledged by the world of learning, perhaps helped by the fact that he had no competition – grammar and linguistics were not major intellectual concerns in European antiquity. Aryabhata's name is honoured in India but, elsewhere, his mathematical work has not yet found the kind of deep scholarly attention and admiration from knowledgeable modern historians – compounded, as we have seen, by a degree of incomprehension – that has been Panini's for a long time. That is beginning to change; his place in the pantheon of mathematical greats seems assured.

[9]H. J. M. Bos' studies of the roots of Leibnizian calculus are very illuminating on this question; see for instance the article "The Fundamental Concepts of Leibnizian Calculus", in his *Lectures in the History of Mathematics*, American Mathematical Society and London Mathematical Society (1991). Newton's approach to foundational issues was more empirical and less logically rigorous.

8

From Brahmagupta to Bhaskara II to Narayana

8.1 Mathematics Moves South

Aryabhata's life work was done by the time he was twentythree. If he wrote other books in later life, they have not come down to us (nor has his supposedly earlier work based on the midnight doctrine of time-keeping). We do not know when he died. One or two names of his purported direct disciples are mentioned by later authors. If they were, it is of no consequence: their work also has been lost.

But within a generation or two after the composition of the *Āryabhaṭīya*, its central ideas were known to Varahamihira in Ujjain at the other end of the north Indian plain, more than a thousand kilometers to the west of Kusumapura. And they had an impact; not only are trigonometric functions to be found all over the independent chapters of the *Pañcasiddhāntikā*, there are also two specific references, once by name and the other time disapprovingly and anonymously, leaving no doubt about who is meant since the object of the disapproval is the hypothesis of the rotating earth. By the time the 6th century came to an end, Indian astronomy had begun to make the transition from the competing models of the named Siddhantas of the *Pañcasiddhāntikā* to a canonically founded science and that was largely Aryabhata's science. But tensions clearly remained at the doctrinal and, sometimes, scientific level: Brahmagupta's questioning of the idea that the sun was much farther away than the moon and Bhaskara I's defence of it, for example. The two were not only close contemporaries, they were also neighbours; they worked in Ujjain or nearby and/or Valabhi (in modern Gujarat) and they both very possibly traced their family origins to Valabhi (which, during this period, was a region where brahminical orthodoxy flourished along with scholarship). The curious fact is that they – one totally loyal to Aryabhata's astronomical ideas but unable to adjust

P. P. Divakaran, *The Mathematics of India*, Sources and Studies in the History of Mathematics and Physical Sciences, https://doi.org/10.1007/978-981-13-1774-3_8

to his new geometry and the other putting up a fight on behalf of orthodoxy, but intellectually equipped to follow his subtleties – seem not to have known each other, going by their writings. Each had his adherents among the succeeding generations but the polarisation never developed into a schism, despite frequent critical references in their writings to those who did not agree with them as *piśāca*, ignoramuses in polite translation. By the middle of the 12th century when Bhaskara II – the universally admired Bhāskarācārya – produced his compendium of the current state of knowledge of astronomy and mathematics, *Līlāvatī*, *Bījagaṇita* and the *Golādhyāya* of *Siddhāntaśiromaṇi*, there was one coherent doctrine founded on the Aryabhatan model (the *Kālakriyā* and the *Gola*), solidly underpinned by the mathematics of the *Gaṇita*. There were also developments which stood somewhat outside the canon, notably Brahmagupta's initiation of the study of the quadratic Diophantine equation and his theorems on cyclic quadrilaterals (the new circle geometry), to the elucidation of both of which Bhaskara II contributed. The last step in the solution of the Diophantine problem was taken before his time but it is his account of the so-called cyclic method (*cakravāla*) that made a difficult subject accessible. The new circle geometry had to wait for the remarkable Narayana, a further two centuries into the future, before it too attained a similar degree of finish. All in all, Bhaskara II marks the beginning of the completion of a programme, a point of closure.

The five centuries separating Bhaskara II from Brahmagupta saw not only a consolidation of astronomical and mathematical knowledge but also an unprecedented geographical spread of scientific activity into regions where no such interest existed earlier, at least not as far as records go. Until the turn of the millennium, northern India was reasonably stable politically, with no more than the usual distractions of the rise and decline of kingdoms and dynasties interrupting public life. There were no wars on a scale big enough to cause wholesale displacements of populations and there was only one armed incursion of any political-military significance from outside, that of a general of the Abbasid Caliphate into the delta of the Indus river (modern Sind, not so far from Valabhi and Avanti) in the beginning of the 8th century. This first Islamic beachhead was the agency by which the sciences of Hind, the works of Brahmagupta among them, reached Baghdad. Politically, it remained localised, the rest of India not seeming to take much notice of it. Life went on as usual. Literature and the arts thrived. New people-friendly forms of Hinduism became popular, symbolised as much by the vast body of religious-mythical literature of the Puranas as by the splendid new temples, of a scale and elaboration never before attempted. And as always during times of stability and prosperity, people travelled, within India and beyond its borders. Aryabhatan astronomy reached China and the Indian zero got engraved on Cambodian and Sumatran stone, a seemingly small thing in comparison with the magnificent Hindu and Buddhist monuments built of the same stone.

Of greater significance for the future than this outreach was that, within India itself, astronomy and mathematics broke out of their confines in the

northern plains and travelled south of the Vindhya ranges to the Deccan plateau and beyond. The circumstances that encouraged the trend are understood in broad terms though all the pieces of the puzzle may not yet find a perfect fit. Politically, the northern and southern halves of India were never under one rule after Ashoka. But Buddhism and Jainism had, from the beginning, found a welcome in many parts of the South and brahmanic learning and lore were not unknown; by the 5th century CE already, there were centres of learning such as the ones in Kanchipuram (in present day Tamilnadu) catering to all three main faiths. Very soon after, there arose powerful new ruling clans and kingdoms, all Hindu, with control over large domains and, even, territorial ambitions across the Vindhya.Their spiritual and cultural gaze was directed to the north; religious practices became uniform and sacred monuments were created in large numbers as though in imitation of the imperial Guptas, thematically and often stylistically indistinguishable from their northern counterparts. There are records of the movement of guilds of architects and sculptors over long distances, going where they found patronage. And, beginning in the 6th century, we have inscriptions testifying to the sponsored migrations of Brahmins from the north – Valabhi being one of the centres of supply – brought in to lend ritual legitimacy to the local rulers, earlier beyond brahminical reach, in return for royal patronage and largesse. Among the priests and the ritualists were also men of knowledge in various disciplines – from Vedic expertise to the science of war – and so it was that Sanskrit, the language of both ritual and learning, came to south India. Eventually, it became such a deep and broad influence on everything to do with language – literary forms, vocabulary, syntax and grammar, even the syllabary – as to turn Telugu, Kannada and Malayalam into virtually new languages. Only Tamil escaped this process of wholesale Sanskritisation (a word I use in its exclusively linguistic sense).

It is against this backdrop of a cultural uniformisation that we have to see the writing of the first mathematical and astronomical books in southern India. That took place in the 9th century under the direct patronage of two different kingdoms, and it is the first time as far as we know that royal support was extended, as acknowledged by the texts, to mathematicians and astronomers. First and better known is Mahavira who was at the court of a Rashtrakuta king Amoghavarsha – the Rashtrakutas were ambitious and powerful; among the monuments they built are some of the great monolithic temples of Ellora – ruling from Manyakheta (modern Malkhed), situated at almost the exact centre of the Deccan. That Mahavira was a Jain while his royal patron was a Hindu seems to have been nothing unusual, very much in the tradition of a righteous king. As far as the mathematics is concerned, his *Gaṇitasārasaṃgraha* is in Sanskrit and in the mainstream (with some minor oddities) of what we may temporarily call north Indian arithmetic and algebra, including series and *kuṭṭaka*. What geometry there is is mensurational and there is no Aryabhatan trigonometry. That is understandable since Mahavira is not concerned with astronomy, no return to the early Jaina speculative cosmogony for him. His mathematical line seems to have died with him except perhaps for an uncertain

connection to Narayana five centuries hence; the Rashtrakutas went into decline and we know of no significant work that can be placed definitely in the area after his passing.

The other seeding took place further south and west, in central and south Kerala in fact, and the seed planted was that of Aryabhatan astronomy as expounded by Bhaskara I. Given the later developments, this early arrival in deepest Kerala of a science that had till then flourished in north India is of great and obvious historical interest. The evidence comes from one book, a commentary in Sanskrit verse and prose on the *Laghubhāskarīya*, the shorter of the two books of Bhaskara I known under his own name, both dealing with the Aryabhatan doctrine (though not a direct commentary like his *Āryabhaṭīyabhāṣya*). The author is Shankaranarayana and the work is known descriptively as *Laghubhāskarīyavivaraṇa* or simply as *Śaṅkaranārāyaṇīya*. It is dated internally to 869 CE and was composed in Mahodayapuram (modern Kodungallur, just north of Kochi), the capital of the so-called second Chera empire and a port city famous in the antique world from Phoenicia to China, from the beginning of the common era, as Musiris or Muziris. The commentary often takes the form of a dialogue between the astronomer royal and his patron the king, questions posed and answered by the two protagonists, and many fine points of astronomical computation come up in the discussion. Among the phenomena that figure in the dialogue is a near-total solar eclipse in 866 CE – which, coincidentally, is also mentioned in a Rashtrakuta inscription of king Amoghavarsha – that can be authenticated by modern astronomical software. (Unfortunately, it is not clear from the text whether the eclipse was actually forecast). There is a reference to what may have been an observatory (*yantravalaya*, a term found also in *Āryabhaṭīya*, a circle of instruments? a circular (spherical) instrument?) and to customs related to time-keeping practised in the capital city such as the beating of drums at regular intervals. Shankaranarayana is a beguiling writer with an interest in life around him, but when it comes to his profession he is fiercely rational: among his sermons in support of astronomy as a science are to be found mocking criticisms of the Puranas and the epics (*itihāsa*) for propagating myths (about eclipses for instance), a side-swipe at Brahmagupta perhaps. The mathematical astronomy in the book confirms him as a competent and sophisticated follower of Aryabhata seen through the eyes of Bhaskara I: the sine table is of the sines, not their differences. Bhaskara I's influence remained strong in Kerala all through, until astronomical research itself came to an end with the fading of the Nila school.

According to a casual allusion by Nilakantha, Shankaranarayana was a disciple of Govinda (Govindasvāmi), himself a commentator of the longer (*Mahā*) *Bhāskarīya*. The connection gets support from the fact that among the earlier astronomers named by Shankaranarayana in his book only Govinda (apart of course from Aryabhata himself) is given the honorific *ācārya*; also, a long passage on *kuṭṭaka* is qualified as *govindakṛti*, composed by Govinda. It is therefore of interest to know whether Govinda himself lived in Kerala and whether he was the carrier of the teachings of Aryabhata and Bhaskara I to the deep southwest

of India. Valabhi was one of the two regions of north India mentioned in legends – and there is some inscriptional support for this – concerning the original domicile of the Brahmins of the west coast. Whether Govinda's origins were in the land of Bhaskara I or not, there is an unambiguous pointer to where he ended up. Commenting on a passage of Bhaskara on the dependence of the rise times of the signs of the zodiac on the latitude of a particular place (*svadeśa*, one's own place), Shankaranarayana gives (in *kaṭapayādi*) the rise times of six of them at a place named Kollapuri, from which we can work out its latitude. Correcting the inevitable copying errors (which, as it happens, can be reliably done), the latitude turns out to be 8.9 degrees. There are other problems in the book in which this latitude is an input, numerically given as 531 minutes. And, to clinch the case, there is a reference in the book to Kollapuri as belonging to the country of his father (*paitryarāṣṭra*).

At 9 degrees latitude, roughly 150 km south of Kodungallur, lies the town of Kollam. It has had that name from at least the beginning of the 9th century when it was already a prosperous port ruled by feudatories of the Chera kings and well known to the traders of the Indian Ocean; Kollapuri, the city of Kollam, is just its partial Sanskritisation. The conclusion is inescapable that Shankaranarayana was a native of Kollam, almost at the southern tip of India, that the beginning of the Aryabhatan tradition of Kerala goes back at least to Govinda in the first half of the 9th century, and that Kollam was its cradle. It may be added that the organised migration of Brahmins into Kerala had begun well before then, the most celebrated from among the Brahmin settlers being the spiritual philosopher Shankara (Śaṅkarācārya, probably 8th century).

As an astronomical text in the Aryabhatan line, *Śaṅkaranārāyaṇīya* is not of major importance. Its value – and the justification of this longish digression into text-analysis[1] – lies in the richness of its historical references, providing not only new insights into the roots of the great breakthroughs of Madhava, 500 years into the future, but also a rare well-documented episode in the story of how the Aryabhatan doctrine spread to remote parts of India.[2]

The state of relative peace in north India came to an end in 1000-1001 CE when the first of many invasions by the sultan Mahmud of Ghazni (in Afghanistan) took place. Over the next twentyfive years an Afghan raid was almost an annual event. Unlike earlier invaders through the northwestern passes, Mahmud was not primarily interested in territory. He came for the loot; even though his main targets were the Hindu temples – grown grand and enormously wealthy in the preceding centuries – and large numbers of people were put to death, it is also not clear that religious zealotry was the dominant motivation.

---

[1]Those who wish to delve into the details will find them in B. Bavare, M. Shetti and P. P. Divakaran, "*Laghubhāskarīya-vivaraṇa* of Śaṅkaranārāyaṇa: its significance in the history of Kerala astronomy", Indian J. Hist. Sci, vol. **45.1** (2010), p.47.

[2]There is a belief in Kerala that the astronomer Haridatta who revised the parameters of Aryabhata in a dated work (684) was a native. There is no textual or inferential evidence pointing either way.

The motive hardly mattered; the outcome was devastation for the western half of the north Indian plain, especially brutal in Valabhi (which had already had a similar experience at the hands of the earlier Islamic rulers of Sind) and Avanti, the land of Bhaskara I and Brahmagupta.

Among those attached to Mahmud's court as an astrologer-astronomer was Alberuni (I follow the spelling adopted by Sachau, the translator of his book on India) who, either as a camp follower or on his own, spent several years in different places in north India, trying to come to grips with Indian civilisation in all its aspects. Later, after Mahmud's death, he wrote up what he learned in the famous book generally known by its shortened title *India*. Alberuni was a competent astronomer and mathematician and devotes a disproportionate amount of space to the practice of those sciences in India (as well as to astrology and related superstition). But the real distinction of the book is that it is the first account we have of Indian science as a whole in its own setting, as seen through the eyes of an outside observer who was also well-versed in Arab and Greek systems of knowledge; and he took pains to get inside a milieu of learning in which he was not at home. The book has much historical value for that reason and for the snippets of revealing information it is full of. Of interest to us (all translations that follow are Sachau's): "This is the reason, too, why Hindu sciences have retired far away from those parts of the country conquered by us and fled to places which our hands cannot yet reach, to Kashmir, Benares and other places" (after saying that "Mahmud utterly ruined the prosperity of the country, ... Hindus became like atoms of dust scattered in all directions"). Or: "Such a disposition (of continents and oceans) of the (spherical) earth is required by the law of gravitation, for according to them (the Hindus) the earth is in the centre of the universe, and everything heavy gravitates towards it". Or, again, speaking of those who criticised Aryabhata's spinning spherical earth: "But Brahmagupta does not agree with them ... apparently because he thought that all heavy things are attracted towards the centre of the earth". (Alberuni goes on to say that there were other reasons why a rotating earth was impossible).

But, as far as mathematics and astronomy are concerned, there is little in *India* that we cannot learn better from earlier Indian texts and much that is to be discounted. A lot of attention is paid to astrology – perhaps naturally – and other popular beliefs, with not always an effort to separate them from the science. This is understandable: if Alberuni limited his travels to the territories ravaged by Mahmud, he is more likely to have met astrologers and almanac-makers than good astronomers. The oft-quoted passage from the opening chapter, "I can only compare their mathematical and astronomical literature, as far as I know it, to a mixture of pearl shells and sour dates, or of pearls and dung. ... Both things are equal in their eyes, since they cannot raise themselves to the methods of a strictly scientific deduction" is not a bad description of Alberuni's own book. The passage is preceded by "... the so-called scientific theorems of the Hindus are in a state of utter confusion, devoid of any logical order, and in the last instance always mixed up with the silly notions of the crowd, e.g. immense

numbers, enormous spaces of time, ..."; so we know what some of the sour dates were: unbounded numbers and infinitely extended time were anathema to Islamic science. To today's reader, more surprising than Alberuni's response to what he found in India is the seriousness with which the book was taken as a guide to the understanding of Indian astronomy (and mathematics and, to a lesser extent, other sciences) from the time Sachau's translation was published in London (1888). Maybe it had something to do with the times; towards the end of the 19th century, the objective scholarship and open-mindedness of the Britain of a century earlier – Jones, Playfair, *et al.* – had been overtaken by an ill-informed and dogmatic colonial ('orientalist') mindset, well-represented by G. R. Kaye's writings on the Bakhshali manuscript and the *Śulbasūtra*.

Among those "scattered like atoms of dust in all directions" were Brahmins in large numbers. Many went south, settling along the way, either temporarily and then moving on, or permanently. We know from the history of Kerala that this was the period when the many waves of a second migration that travelled gradually along the western coast and eventually ended up there started. It continued till about 1300 or 1400; according to the estimate of one eminent historian, the proportion of Brahmins in the population of Kerala may have been as high as 20-25% in the 14th century. This was also the time when the Sanskritisation of the language of Kerala from a local dialect of the Tamil of the time to what is recognisably Malayalam had been completed. The brahminical sciences, *śāstra*, came with the Sanskrit; in particular, the new arrivals brought fresh oxygen to the almost dead astronomical embers of Kollam and Mahodayapuram. The times match: Madhava was born around 1350 and his caste name, Emprantiri, identifies his family as having recently arrived in Kerala from further north. That story will occupy us more fully in a later chapter.

For now, we return briefly to the North. Mahmud was followed a hundred and fifty years later by another Afghan sultan, Muhammad Ghori. Ghori combined territorial and political ambition with greed for treasure and religious zeal. As every schoolchild in India knows, his foray into India started the long process that led on to the domination of most of north (and, intermittently, parts of south) India by Afghan and Turkic dynasties and, eventually, to Babar and the founding of the Mughal (Mongol) empireMughal empire. The aftermath of the Ghori invasion was like nothing that India had experienced before. It overwhelmed large parts of northern India, all the way to Bihar and Bengal to the east, and consciously targeted the ideal of cultural plurality that had sustained India – even if sometimes only in lip service – through good times and bad; henceforth, in the areas of Turkic domination, there was to be only one official ideology and all that stood against it was to be annihilated. Universities and their libraries, repositories of all that is permanent in a civilisation, were favoured targets, not for the first time in history nor the last. Within ten years of his arrival in India proper, a Turkish general of Ghori named Bakhtiyar took it on as his holy duty to pillage and burn down Nalanda and the other universities of eastern In-

dia – themselves risen from the ashes of Taxila, long ago and far away – putting to the sword the monks and teachers who could not or did not wish to escape.[3]

It is in the interregnum between Ghazni and Ghori that Bhaskara II lived his life. We know from his own writings that he was born in 1114 and that his masterpieces were composed in the middle of the century. This is supplemented by a unique (as far as I know) secondary source of information about his ancestry, a long and remarkably informative inscription on a plaque installed by his own grandson in a religious retreat (*maṭha*) and school in a village near the modern town of Chalisgaon in Maharashtra, about 300 km northeast of Mumbai. Firstly, the circumstances suggest that the location was his natal village, at least the home village of his family. Chalisgaon is only about the same distance from Ujjain as it is from Mumbai but, significantly, it is to the south of the Vindhya mountains which Ghazni did not cross. At the time of the Ghazni raids, Malava, to the north of the Vindhya, was ruled by king Bhoja who was involved in the resistance to the Afghans and was also an author himself and a patron of learning. The plaque mentions that a direct paternal ancestor of Bhaskara five generations earlier had received honours from Bhoja. (For the full inscription together with a French translation, see Patte [SiSi-P]). It is therefore entirely possible that the family was originally from Malava – Bhoja's capital was Dhar, close to Ujjain – and that it moved south as a result of the troubles. That would make Bhaskara II yet another of the migrant mathematicians, though his ancestors did not have to cover anywhere near as great a distance as Aryabhata's or Madhava's.

There was little mathematics of true significance from north India after the Ghori invasion, at least none of unquestioned northern provenance, not until the 16th-17th century when, under Akbar, the Moghul empire settled down to stability and peace. But memory and the traditional system of passing on knowledge must have lingered on. Colebrooke writes that when he was trying to establish the chronology of the famous astronomers of the past, it was among the pandits of Ujjain that a collaborator of his found help (at the end of the 18th century); and, as far as Brahmagupta and Bhaskara II were concerned, the information was accurate. The fact remains, nevertheless, that mathematics and astronomy had moved south, a move which went unremarked for a long time. In the preface to his edition of *Āryabhaṭīya* Kern (in 1874) speaks of the time ".... after the great Bhaskara, in an age when the living breath of science had already parted from India ... ". This was understandable because very few people in Europe at that time had any idea of the Nila rejuvenation, even though Kern himself had in his possession Charles Whish's collection of manuscripts from Kerala and his edition was itself based on Parameshvara's commentary.

[3] Unlike in the story of Taxila, there is plenty of chillingly graphic documentation from contemporary Islamic chroniclers on what happened. For a sample, see the small collection of quotations in the last chapter (on the great universities) of Warder's book, A. K. Warder, *Indian Buddhism* (Third revised edition, 2000), Motilal Banarsidass, New Delhi.

8.2 The Quadratic Diophantine Problem – *bhāvanā*

Of the two mathematical themes, one algebraic (number-theoretic) and the other geometric, connecting Bhaskara II with Brahmagupta, the stronger link is the circle of ideas around the solution (and the general properties that lead to the solution) of the equation called *vargaprakṛti* ('of square nature', more simply, quadratic; Brahmagupta himself does not employ the term). The other theme, the new circle geometry concerned with cyclic quadrilaterals, will be taken up in the next but one section.

The problem in its general form (which I will call from now on the quadratic Diophantine problem or simply the quadratic problem) is to solve the equation (as it is generally written)

$$Nx^2 + C = y^2$$

for x and y in integers, where the coefficient ('the multiplier') N is any non-square positive integer and C ('the interpolator') is any integer, positive or negative. It occurs, in this general algebraic form and not merely as a numerical problem, in Chapter XVIII (called *Kuṭṭakādhyāya*; a more accurate and modern description would be as the 'chapter on number theory') of the *Brāhmasphuṭasiddhānta*. The problem was invented by Brahmagupta. No Indian text before him has anything similar. From outside India, one version of Archimedes' 'cattle problem' leads to an equation of this type with specified numerical coefficients; its solution comes out as absurdly large numbers (unlikely to have been worked out by Archimedes, according to Heath in his edition of the works of Archimedes). The equation had no natural application to astronomy though Brahmagupta throws in some astronomical terminology in the numerical examples he uses to illustrate the general principles. Chapter XVIII has a few other Diophantine equations, equally 'pure', all more or less inspired, it would seem, by the challenges illustrated by Arayabhata's much simpler linear problem. Indeed, Brahmagupta concludes the chapter by saying that such questions are stated for gratification (Colebrooke's translation of *sukha*, [Co]). That he himself thought of these questions as enjoyable mathematics and was proud of what he achieved is clear from the illustrative problems he poses throughout the chapter, ending them often with a declaration like "The person who can solve this problem is versed in *kuṭṭaka*". The bar is raised when it comes to quadratic problems: "The person who can solve this problem within a year is a mathematician", somewhat reminiscent of the challenge accompanying Archimedes' cattle problem.

For any N and C, if (x, y) is a solution of the equation, so are $(-x, y), (x, -y)$ and $(-x, -y)$. Indian authors generally look only at solutions in which both x and y are positive; from now on, the term 'integral solution' (or, sometimes, just 'solution') will mean a positive integral solution unless otherwise qualified. On the set of all such solutions (for fixed N and C) there is an order: if (x, y) and (x', y') are solutions, $N(x^2 - x'^2) = y^2 - y'^2$, and hence $x > x'$ if and only if $y > y'$. In particular there is a minimal solution (X, Y).

(Later in Europe, the minimal solution for $C = 1$ was recognised as the fundamental solution from which all others can be generated. It is not to be confused terminologically with the 'least root' employed by some to translate the Sanskrit technical term for the value of x in a solution; on the occasions when the values of x and y need to be separately denoted, I will use the terms 'first root' and 'second root'). Some easy-to-see conditions on N and C are also best dealt with here. A solution of the form $(0, y)$ exists only if C is a square, and of the form $(x, 0)$ only if C is negative and of the form $C = -N\times$ (a square). If $C = 0$, rational solutions exist only if N is a square and the integral solutions are of the form $(x, \sqrt{N}x)$, x any integer. For $C = \pm 1$ and N a square, from $y^2 - Nx^2 = \pm 1$ we see that the only solution is $x = 0, y = 1$ or $x = 1, y = 0$ (with $N = 1$) for the two signs of C. In fact, when N is a square, the equation has only a finite number of solutions (including the possibility of no solutions) for any C, as we can see by enumeration of $C = \pm 1, \pm 2, \cdots$. More generally, there are values of N, not a square, and C (small values can again be found by enumeration) for which there are no solutions.

Brahmagupta did not give a method for finding all solutions of a given quadratic Diophantine equation for given values of N and C. What he did was deeper than setting down an algorithm; he created a theory of such equations which he then used to find solutions of some equations and which eventually became the essential input in the later work on the general solution. In this respect, the historical sequence is to an extent the opposite of what transpired in the study of the linear Diophantine problem. In that instance, the first breakthrough was Aryabhata's algorithm and the general theoretical framework was built up around it over a period of time (Chapter 7.2). In the case of the quadratic problem, the structural foundation put in place by Brahmagupta hardly needed any improvement – later authors, by and large, merely made explicit what is conveyed implicitly and through examples by him – until the final step of finding all (with some reservation, see later) solutions. That step was taken by Jayadeva as we know from an extract in a commentary on the *Laghubhāskarīya* by Udayadivākara, dated 1073 (i.e., after the Ghazni invasions; some historians have suggested that he hailed from Kerala). Nothing else is known about Jayadeva but, at least, it is certain that the quadratic problem was sorted out by the 11th century. (The chronological aspect is actually quite intriguing. Śrīpati, a noted commentator on Brahmagupta's number theory, does not have the general solution and he lived in the first half of the 11th century. So, either Jayadeva worked in the third quarter of the 11th century or, less likely, earlier in some remote place beyond the horizon). It is in fact through Bhaskara II's account in *Bījagaṇita* (which does not mention Jayadeva) that the world knew, until recently, about the details of the solution (Datta and Singh [DS], for example, do not know Jayadeva and attribute the solution to Bhaskara)[4].

[4]The quotations from Jayadeva in Udayadivakara's *Sundarī* were first brought to the notice of modern scholars in K. S. Shukla's "Ācārya Jayadeva, the Mathematician", Gaṇita **5** (1954), p.1. For historians of Indian arithmetic, Shukla's discovery was obviously a milestone.

The central idea of Brahmagupta is to investigate what we now call the space of solutions of the equation without actually having an explicit solution; i.e., to study the properties that the set of solutions must have following from the fact that they all satisfy the equation and to ask to what extent these properties determine the solutions. In the explicit, exact, form in which he implemented the idea and in its subtlety and power, it was a first, but the general philosophy of getting information about solutions from the structure of an equation goes back to earlier times, for example the algebraic Bakhshali square root formula: to find the square root of n, choose an arbitrary number m, determine the equation satisfied by $\sqrt{n}-m$ from the equation for $\sqrt{n}$, and so on (Chapter 6.5). Indeed, a solution pair (x, y) is referred to as 'roots' (*mūla*) with various qualifiers and the choice of an arbitrary (*iṣṭa* or a synonymous word, translated as 'assumed' (Colebrooke) or 'optional' (Datta and Singh)) number is often the starting point in a search for the roots.

The philosophy comes through clearly already in the special case $C = 1$. Accordingly, consider the Diophantine equation

$$Nx^2 + 1 = y^2,$$

which I will call the special (quadratic) equation; it is special also in that it was the main object of interest, with the general problem with $C \neq 1$ playing a supporting role. If $(x, y)_1 = (x_1, y_1)$ and $(x, y)_2$ are solutions of it (for a fixed N), Brahmagupta observes that they can be composed by a process involving cross and direct multiplication according to the rule (the reason for the apparently clumsy notation $\times_+$ will become clear in a minute)

$$(x, y)_1 \times_+ (x, y)_2 = (x_1y_2 + x_2y_1, Nx_1x_2 + y_1y_2)$$

to obtain a new solution of the same equation. (The technical term for the composition rule, not used by Brahmagupta himself, is *bhāvanā*, derived from the root for 'to make exist' or 'create'). The proof of the assertion is elementary. Start from the algebraic identity (sometimes called Brahmagupta's (additive) identity)

$$(y_1y_2 + Nx_1x_2)^2 - N(x_1y_2 + x_2y_1)^2 = (y_1^2 - Nx_1^2)(y_2^2 - Nx_2^2)$$

which holds for any x_1, x_2, y_1, y_2 and N, not necessarily integers. If $(x, y)_1$ and $(x, y)_2$ solve the equation, the two factors on the right are both unity and the result follows. Brahmagupta does not provide a proof. Later authors do and it is essentially what is given above.

We can compose a solution with itself: if (x, y) solves the special equation, so does

$$(x, y) \times_+ (x, y) = (2xy, y^2 + Nx^2).$$

An immediate corollary is that when N is not a square, the equation has an infinity of solutions if it has one, unbounded above (looking at only positive solutions), produced by repeated composition; in particular, the minimal solution will generate an infinite sequence of ever bigger solutions. If on the other

hand N is a square, the unique solution is $(0, 1)$ as noted earlier in this section, and its *bhāvanā* with itself is itself.

The rule of composition has two easy generalisations. Firstly, the additive identity of Brahmagupta has a subtractive counterpart:

$$(y_1y_2 - Nx_1x_2)^2 - N(x_1y_2 - x_2y_1)^2 = (y_1^2 - Nx_1^2)(y_2^2 - Nx_2^2).$$

Correspondingly, there is also a subtractive composition of solutions:

$$(x, y)_1 \times_- (x, y)_2 = (x_1y_2 - x_2y_1, y_1y_2 - Nx_1x_2).$$

Subtractive composition is just the additive composition with one root in one of the factors reversed in sign, say $(x, y)_1 \to (-x, y)_1$, $(x, y)_2 \to (x, y)_2$. The other reversals of sign give nothing new. Subtractive composition is of interest because we can get new positive solutions by so composing two positive solutions, different in principle from their additive composite.

The other generalisation is just as easy to establish but conceptually deeper and more productive in consequences. Brahmagupta's identities do not involve the interpolator C. If $(x, y)_i, i = 1, 2$, are now solutions of the general equation with interpolators C_i respectively, the right side of Brahmagupta's identities is just the product C_1C_2. So the composition law in full generality is:

If $(x, y)_i$ is a solution of the general quadratic Diophantine equation

$$Nx^2 + C_i = y^2,$$

then

$$(x, y)_1 \times_\pm (x, y)_2 = (x_1y_2 \pm x_2y_1, y_1y_2 \pm Nx_1x_2)$$

are solutions of the equation

$$Nx^2 + C_1C_2 = y^2.$$

It is in this general form that Brahmagupta formulates, in the first two verses of the relevant section of Chapter XVIII, the principle of *bhāvanā*. The wording leaves no room for doubt that the interpolators are taken to be any two integers and that the solutions being composed are also arbitrary solutions. Two special cases are of great practical utility: i) by composing two solutions $(x, y)_i$ $(i = 1, 2)$ with the same interpolator C, we get solutions $(x_1y_2 \pm x_2y_1, Nx_1x_2 \pm y_1y_2)$ with interpolator C^2; ii) by composing repeatedly a solution (x, y) with interpolator C with the solution (x', y') of the special equation with interpolator 1 (or by composing (x, y) with the infinity of solutions of the special equation) we get an infinite number of solutions with interpolator C.

The rest of Brahmagupta's section on the subject augments *bhāvanā* with a few other general statements and illustrates them by means of several numerical examples.

The key conceptual advance in the passage to general composition was thus to look at the family of equations parametrised by C, and their solutions,

in one go rather than at individual equations for each value of C separately. In modern language, we would say that the object Brahmagupta studies is the space Σ of all solutions for all values of C; an element of Σ is a triple $(x, y; C)$ of integers and we would think of *bhāvanā* as a binary composition in Σ, i.e., as (commutative) maps $\Sigma \times \Sigma \to \Sigma$ given by the modified products

$$(x, y; C) \times_{\pm} (x', y'; C') = (xy' \pm x'y, yy' \pm Nxx'; CC').$$

Though everything was expressed in words, fundamentally this seems to be how Brahmagupta's own mind ran; the generality of the statement of the principle and the fact that it is the first proposition on the subject in the *Brāhmasphuṭasiddhānta* argue for it. Given the level of sophistication of his algebra elsewhere – he was after all the first to lay the base for formal algebraic reasoning – we can believe that he was at home in the degree of abstraction demanded.

Whatever the reason why Brahmagupta chose to investigate these questions and however he thought his way through them, it is a fact that his treatment of the quadratic Diophantine equation, specifically the principle of *bhāvanā*, has been the part of Indian mathematics to have attracted the greatest interest among knowledgeable modern mathematicians (I am not speaking here of professional historians of mathematics), definitely greater than in Aryabhata's functional differences and the discrete fundamental theorem of calculus. The main reason for this is of course that it fits well into the metaphysics of modern number theory in the way it combines an abstract and structural viewpoint, easy to grasp once stated, with ingenious technical virtuosity. For number-theorists it seems to have been a 'natural', a highly nontrivial but not impossibly difficult problem. It is well known that when Fermat, generally considered to be the founder of modern number theory, wanted to interest mid-17th century Europe in the richness of arithmetic, one of the questions he challenged his contemporaries with was the solution of the special quadratic Diophantine equation – what is it about the problem that provokes its devotees to issue these boastful challenges, whether in ancient Greece or India or Europe? – which became famous in course of time as, for no good reason, Pell's equation. About a century after Fermat, Euler and Lagrange both worked out the theory of such equations, integrating into it the properties of continued fractions and discovering along the way the composition law and its crucial role in the general theory. Still later, it became part of the study of new algebraic objects generalising the notion of integers. The subject is still very much in the spotlight.[5]

In India, *vargaprakṛti* remained an object of intense study and research for a long time. Many fine mathematicians contributed to the elucidation of prin-

[5]For an account of the history of the quadratic Diophantine equation in the context of the modern theory of numbers, see André WeilWeil, André, *Number Theory: An approach through history from Hammurapi to Legendre*, Birkhauser, Boston (1984) (but disregard his scepticism about Indian proofs). Amartya Dutta's article cited earlier as well as his "Brahmagupta's Bhāvanā: Some Reflections", in [CHIM] have quotations and references which will bring the reader up to date. The topic and its extensions to higher degrees remain a vibrant field of research in modern algebraic number theory.

ciples and methods that were explicitly stated as rules, or could be extracted from the examples, in Brahmagupta's own work. There were some who broke new ground and among them belongs Jayadeva who invented a systematic algorithm (to which he gave the name *cakravāla*, the 'cyclic algorithm') for the solution of the special quadratic equation by integrating it with the theory of linear equations which could be handled by *kuṭṭaka*. Bhaskara's account of the method in *Bījagaṇita*, illustrated with nontrivial numerical examples, comes less than a hundred years after Jayadeva. The account below of the development of the subject is from the vantage point of what Bhaskara knew, supported by Jayadeva's own original words. The other innovator, Narayana, lived in the 14th century and his basic idea was to link the infinity of increasing solutions (x_n, y_n) of the equation $Nx^2+1=y^2$ to better and better rational approximations of $\sqrt{N}$ obtained by neglecting 1 in comparison with Nx^2: $\sqrt{N}=y_n/x_n$. Rational approximations of square roots is a topic tailor-made for the use of continued fraction methods; Narayana put them to good use and thus his work foreshadows the European approach to Pell's equation.

This aspect of Narayana's work will not be further discussed here (nor in this book) except to note that numerical determination of square roots is unlikely to have been his (or Brahmagupta's initial) main motivation for studying *vargaprakṛti* – there was already the excellent and user-friendly Bakhshali algorithm for that purpose. Somewhat speculatively, I would suggest that Narayana's prime interest was more theoretical (as it probably was also of Brahmagupta, see the next section), namely the proposition: given an integer N not a square, there is an infinite sequence of (increasing) positive integers x such that Nx^2+1 is a square with, as a corollary, the implication that $\sqrt{N}$ can be approximated arbitrarily closely (*āsanna*, to import a term from Aryabhata and Nilakantha) by rationals. The wording of many of the commentators ("the square of the first root multiplied by the multiplier and increased by the interpolator gives the square of the second root") is consistent with such an interpretation. Later on we shall see that other areas of Narayana's work, geometry in particular and combinatorics, had a big impact on the Nila school which, in chronological terms, was contiguous with his life. Madhava's birth fell in Narayana's most productive phase (mid-14th century) and irrationality questions were a concern for the Nila mathematicians, certainly for Nilakantha; it is not very likely that he made his conjecture on the irrationality of π as a casual remark, without deep reflection.

8.3 Methods of Solution – *cakravāla*

Brahmagupta does not have a general algorithm for solving the quadratic Diophantine equation for given values of N and C. Special methods for finding one solution (and hence infinitely many of them), adapted to N and C with particular properties, are what he describes. Some but not all of them are based on the ingenious deployment of the principle of composition. To that end, several

verses following the two on *bhāvanā* are devoted to its corollaries or auxiliary results to be used in conjunction with it, supplemented by illustrative numerical problems. Since there is no general algorithm until we come to *cakravāla*, a look at these examples is useful for an appreciation of the style of Brahmagupta's own approach to the business of actually finding solutions. For that reason (and unlike in my account of the linear problem) a few such numerical problems will be taken up as we go along. It is also perhaps necessary for the present-day reader to remember that the idea of finding all solutions of the equation, or of asking whether a particular procedure will lead to all solutions, is not present in the *Brāhmasphuṭasiddhānta*, nor in later texts. All of them are content with noting that solutions are *ananta*, without end.

Though the statement of the principle of composition covers the general case, it would seem from the tenor of the discussion – the rules as well as the examples – that the primary goal was the finding of one integral solution of the special equation ($C = 1$) which could then be used to generate other solutions of itself; a solution of the general equation ($C \neq 1$, with emphasis on certain special values, see below) seems to have been of secondary interest. It may well be that $C = 1$ is the problem Brahmagupta started out with and that equations with $C \neq 1$ were brought in to facilitate its study. The other preliminary remark is that it was realised that fractional solutions are a good starting point in the search for integral solutions. As has already been noted, composition holds for them just as well as it does for integral solutions and can be used in favourable cases (along with other methods) to go from the former to the latter. But, again, no general procedure for the specialisation is given.

Here are two of Brahmagupta's auxiliary rules, neither relying on *bhāvanā*:

1. Suppose C is a multiple of a square, $C = k^2M$. Then, if (u, v) is a rational solution of $Nx^2 + C = y^2$, $(u/k, v/k)$ is a rational solution of $Nx^2 + M = y^2$ and conversely. In particular, if $C = k^2$ is a square, its rational solutions are related to those of the special equation through scaling by k. Square interpolators are of interest since they can be produced by self-*bhāvanā*.

Brahmagupta gives two examples of the application of the rule, both very elementary. The slightly less trivial is the equation $3x^2 - 800 = y^2$. If (ξ, η) is a solution, then, dividing through by 400, we see that $(\xi/20, \eta/20)$ is a solution of $3x^2 - 2 = y^2$. From the solution (1,1) of the latter, it follows that (20, 20) is an integral solution of the original equation. One suspects that these very easy examples are primarily meant to illustrate the principles involved (though they are also accompanied by the formulaic challenge to 'the mathematician' to solve them within the year) – one might just as easily have solved the original equation 'by inspection' as the derived equation.

2. Suppose N is a square, $N = m^2$. Then for any k,

$$x = \frac{1}{2m}\left(\frac{C}{k} - k\right), \qquad y = \frac{1}{2}\left(\frac{C}{k} + k\right)$$

is a solution of $Nx^2 + C = y^2$ as follows from the identity $(C/k - k)^2 + 4C = (C/k + k)^2$. In particular, if C factorises as $C = kl$, then we have the almost

integral solution $x = (l-k)/(2m), y = (l+k)/2$ (switching k and l only switches the sign of x): if the factors of C can be so chosen that $2m$ divides $l-k$ ($l+k$ then being even), the solution will be integral.

Of the two examples given by Brahmagupta, consider $4x^2 - 60 = y^2$. Among the distinct factorisations of 60, only 30×2 is such that $l - k = 28$ is divisible by $2\sqrt{N} = 4$, giving the integral solution (8, 14).

Later authors gave a variety of rules for rational solutions of the special equation. In one example, the identity (for arbitrary rational v though in practice v is generally an integer) $4Nv^2 + (v^2 - N)^2 = (v^2 + N)^2$ says that

$$x = \frac{2v}{v^2 - N}, \qquad y = \frac{v^2 + N}{v^2 - N}$$

is a 1-parameter family of solutions of $Nx^2 + 1 = y^2$. Interestingly, the same set of rational solutions can also be obtained by *bhāvanā* by taking the trivial identity $N + v^2 - N = v^2$ as saying that $(1, v)$ solves the equation with multiplier N and interpolator $v^2 - N$ (an observation which was to become a key input in *cakravāla*); compose it with itself additively: $(1, v; v^2 - N) \times_+ (1, v; v^2 - N) = (2v, N + v^2; (v^2 - N)^2)$ and divide through by $(v^2 - N)^2$. Still more generally, any pair (u, v) solves the equation with $v^2 - Nu^2$ as the interpolator; composition followed by division by the square of the interpolator leads to the 2-parameter family of rational solutions

$$x = \frac{2uv}{v^2 - Nu^2} \qquad y = \frac{v^2 + Nu^2}{v^2 - Nu^2}$$

of the special equation. This solution also can be easily seen to arise from the quadratic identity in u^2 and v^2, $4Nu^2v^2 + (v^2 - Nu^2)^2 = (v^2 + Nu^2)^2$; all of these are in fact no more than reexpressions of the general identity $(x - y)^2 + 4xy = (x + y)^2$ whose geometrical realisation goes back to the *Śulbasūtra*.

Indeed, the conclusion that solutions constructed by *bhāvanā* can equivalently be stated as consequences of elementary algebraic identities is very general, higher degree *bhāvanā* correspondingly being equivalent to the validity of cubic, etc. identities. The originality of *bhāvanā* is in the fact that, for rational solutions of the special equation, it captures all such identities in one succinct and powerful algebraic principle, to be used recursively to generate an infinity of them – the first serious instance of the use of pure algebra to transcend the dimensional limitations of geometric imagination.

Before turning to the question of integral solutions, it is useful to remind ourselves that Chapter XVIII of the *Brāhmasphuṭasiddhānta*, even the section of it dealing with *vargaprakṛti* (*the* quadratic Diophantine problem in my nomenclature), has a variety of other Diophantine problems some of which are related to it, some not; in our admiration for the theoretical elegance of *bhāvanā*, we sometimes tend not to do them full justice (the book of Datta and Singh ([DS]) being an exception). As an example of one such (linear-quadratic) problem, Brahmagupta asks for x, y and z such that $ax + 1 = y^2$ and $bx + 1 = z^2$, presumably in integers for integral a and b and gives the answer as

$x = 8(a+b)/(a-b)^2, y = (3a+b)/(a-b), z = (a+3b)/(a-b)$. (Integrality is not guaranteed for general a and b but the example he gives, $a = 13, b = 17$, is designed to produce an integral solution). Brahmagupta does not of course tell us how he arrived at the solution, but Bhaskara II does (in fact he describes more than one way of doing so). The method is an interesting mix of elementary algebra and *bhāvanā* and so worth a look as a typical illustration of the algebraic culture that was the setting for Brahmagupta's number theory.

The equations have the trivial solution $x = 0, y = z = 1$. To find other solutions, eliminate the linear term: $az^2 + b - a = by^2$, and multiply through by b, resulting in the *vargaprakṛti* equation $abz^2 + b(b-a) = (by)^2$ (alternatively, divide by b and work with a fractional multiplier and interpolator), with the trivial roots $z = 1$ and $by = b$. We have just seen how to construct a very large family of rational solutions of the special equation; compose this family (with multiplier ab) with the trivial solution for the same multiplier but interpolator $b(b-a)$ to obtain an equally large family of rational solutions (z, by) of the latter, and hence of solutions (x, y, z) of Brahamagupta's original linear-quadratic problem. The choice $u = 1$, $v = a$ in the parametrisation of the general rational solutions described earlier corresponds to Brahmagupta's formulae for x, y and z.

Bhaskara II actually deals with the more general linear-quadratic problem in which the 'interpolators' 1 are replaced by general integers, by essentially the same method as described above. (Datta and Singh ([DS]) observe that this is also the method employed by Lagrange in his 'Additions' to Euler's work on 'Pell's' equation). In fact, by the time of Bhaskara, the investigation of several types of Diophantine problems had become a staple of Indian arithmetic and he himself describes the solutions of many of them through rules and examples. Not all of them are challenging but there are also some hard nuts among them. Mathematically, what is interesting is the rich variety of questions asked and answered, simultaneous equations involving more than two variables, polynomials of degree higher than 2 (and, in one case, of arbitrary degree), etc. etc. Subsequently, the Diophantine spirit was actively cultivated by Narayana and, still later, by Chitrabhanu of the Nila school. (A very thorough account of the work along these lines can be found in [DS]). As interesting for the historian is the fact that it is as an introduction to this more general Diophantine theory of the final chapters of the *Bījagaṇita*, well after he has dealt with *kuṭṭaka*, *bhāvanā* and *cakravāla*, that Bhaskara gives us his poetic endorsement of the power of algebra: as the sun illuminates the lotus, so have intelligent mathematicians of the past, by the use of syllables for the unknowns, made *Bījagaṇita* clear to the dull-witted.

Brahmagupta himself seems to have thought of algebra, Diophantine arithmetic and – given his work on rational and integral triangles and cyclic quadrilaterals (Chapter XII of the *Brāhmasphuṭasiddhānta*, to be described in the next section) – even geometry as different facets of one cogent mathematical theme driven by the concept of numbers, each feeding into the other. In this respect, his geometry is close in spirit to the *Śulbasūtra* – among his results on

the right triangle or, rather, the isosceles triangle with its altitude, is the general formula for all primitive integral Pythagorean triples. It is quite distinct in inspiration from Aryabhata's use of the diagonal theorem, first as the foundation of trigonometry and then as the first step in the eventual development of calculus on the circle: infinitesimal rather than integral geometry.

To return to the question of finding integral solutions of *vargaprakṛti*, from rational solutions or otherwise, it is a fair comment that it remained a matter of ingenuity and art rather than a systematic logical process, for Brahmagupta as much as for those who came after, until the advent of *cakravāla*. One guiding principle was that, in favourable cases, *bhāvanā* can turn fractional solutions into integral ones. The simplest illustration of how this worked is provided by the composition law for the special equation: if (x, y) solves it, so does

$$(x, y)^2_+ = (x, y) \times_+ (x, y) = (2xy, Nx^2 + y^2) = (2xy, 2y^2 - 1)$$

where in the last step N has been eliminated using the original equation satisfied by (x, y). The composite solution is integral if y is integral and x is half-integral. Brahmagupta illustrates this with a pretty example. To solve $92x^2 + 1 = y^2$ in integers, start with the observation that $92 \times 1^2 + 8 = 10^2$. Composing the solution $(1, 10)$ of $92x^2 + 8 = y^2$ with itself, we get the solution $(20, 192)$ of $92x^2 + 64 = y^2$ and, dividing by 64, the fractional solution (5/2, 24) of $92x^2 + 1 = y^2$. Finally, compose this last solution of the special equation with itself, resulting in its integral solution (120, 1151), which can then be used to generate an infinity of integral solutions by further composition.

Composition of higher degree can similarly be used to produce integral solutions from (some) fractional solutions. A helpful fact in doing this is that in the composition of (x, y) with itself any number of times and for any interpolator C, N is always multiplied by x^2 (and other factors) and hence can be eliminated in favour of C by virtue of the original equation $Nx^2 = y^2 - C$ without introducing denominators, i.e., composite solutions can be written as polynomials in x and y with integer coefficients. Thus, for $C = 1$, the results of 3rd and 4th degree *bhāvanā* on the special equation are, in an obvious notation,

$$(x, y)^3_+ = (4xy^2 - x, 4y^3 - 3y)$$

and

$$(x, y)^4_+ = (8xy^3 - 4xy, 8y^4 - 8y^2 - 1),$$

showing that the pairs on the right sides solve the special equation if (x, y) does, in particular that if y is integral and x is quarter-integral (among other possibilities), 4th degree composition leads to an integral solution.

This strategy can be adapted to certain particular values of $C \neq 1$ and, when combined with Brahmagupta's first auxiliary rule for square interpolators (see the beginning of this section), leads to new ways of generating integral solutions of the special equation. The special values investigated are $C = -1, \pm 2, \pm 4$. These particular cases were described systematically only in the first half of the 11th century by Sripati, but the results are already

present in the *Brāhmasphuṭasiddhānta*, either explicitly as rules or in the problems. The first special case says that any integral solution for $C = -1$ leads, by *bhāvanā*, to an integral solution for $C = 1$. The construction is trivial: if (x, y) is a solution for $C = -1$, compose the triple $(x, y; -1)$ with itself, $(x, y; -1)^2_+ = (2xy, Nx^2 + y^2; 1)$. Slightly less trivial are the cases $C = \pm 2$. We have $(x, y; \pm 2)^2_+ = (2xy, 2y^2 \mp 2; 4)$ where, as before, N has been eliminated in favour of C. Dividing through by 4, we see that $(xy, y^2 \mp 1)$ is an integral solution for $C = 1$ if (x, y) is an integral solution for $C = 2$ or $C = -2$ respectively.

The formulae for $C = \pm 4$ are explicitly given by Brahmagupta. These cases are more subtle, $C = 4$ being somewhat simpler to deal with since it is already a square. Suppose the integral pair (x, y) solves $Nx^2 + 4 = y^2$. Then $(x/2, y/2)$ is a rational solution of $Nx^2 + 1 = y^2$. Take its 2nd and 3rd powers:

$$\left(\frac{x}{2}, \frac{y}{2}\right)^2_+ = \left(\frac{xy}{2}, \frac{y^2}{2} - 1\right),$$

$$\left(\frac{x}{2}, \frac{y}{2}\right)^3_+ = \left(\frac{x}{2}(y^2 - 1), \frac{y}{2}(y^2 - 3)\right).$$

The 2nd power is integral if and only if y is even and the 3rd power is integral if y is odd but also if both x and y are even; between them, the two formulae cover all possibilities even though, when x is odd and y is even, the 3rd power formula will not give an integral solution of the special equation. Curiously, Brahmagupta's verse mentions only the 3rd power solution and he illustrates it with an equation where it works, $3x^2 + 4 = y^2$ which is solved by $(2, 4)$. Was it an oversight or did he think it unnecessary to go over a case which yields to just one application of *bhāvanā*?

The way to proceed in the case $C = -4$ is now clear. First divide by 4, then square the resulting rational solution of the equation for $C = -1$ to get to a rational solution for $C = 1$, then do *bhāvanā* again as many times as necessary to get to an integral solution – from the $C = 4$ case one would guess, correctly, that thrice will be enough. (Alternatively, compose $(x, y; -4)$ with itself to go to the equation with $C = 16$ and divide by 4 to reduce it to the previous case $C = 4$). The result (those who wish to know the details will find them in [DS]) is that if (x, y) solves $Nx^2 - 4 = y^2$, then

$$\left(\frac{1}{2}xy(y^2 + 1)(y^2 + 3), \frac{1}{2}(y^2 + 1)(y^2 + 2)(y^2 + 3) - (y^2 + 2)\right)$$

solves $Nx^2 + 1 = y^2$. The solution is clearly integral if x and y are integral. Brahmagupta has a numerical example: $(1, 3)$ solves $13x^2 - 4 = y^2$; the formula then says that $(180, 649)$ solves $13x^2 + 1 = y^2$.

The general conclusion we can draw from all this work is that Brahmagupta's discovery of the principle of *bhāvanā* launched a long period of exploration, beginning with himself, of how it could be made use of in finding

integral solutions of the special equation. The methods, for all their ingenuity, remained *ad hoc* with no unifying thread running through them; there is no indication in the texts that the semi-empirical knowledge so gathered led to any progress of a general nature beyond what Brahmagupta already knew. The running theme remained the construction of auxiliary rational solutions, easy to do, without there being an accompanying investigation of the general conditions under which they could be turned into integral ones. In the event, when a general algorithm, *cakravāla*, was eventually found, it owed little to the rational solutions of the earlier work except in a somewhat expedient and dispensable reliance on the passage from $C = -1, \pm 2, \pm 4$ to $C = 1$, described above. Instead, it took its inspiration directly from the integral arithmetic of the *kuṭṭaka* method of solving the linear Diophantine problem.

As noted already, the algorithm is associated with the name of Jayadeva on the strength of Udayadivakara's attribution and citation of the relevant verses. The term *cakravāla* is Jayadeva's own. The name stuck though it is far from clear what is particularly cyclic or circular about the procedure. Udayadivakara's citation is dated to 1073 but Sripati, whose work in the second quarter of the 11th century contributed significantly to the understanding of rational solutions and to the systematisation of the particular cases $C = -1, \pm 2, \pm 4$, has no knowledge of the algorithm nor, of course, of its name. We can therefore conclude with some confidence that *cakravāla* was invented in the two or three decades separating Udayadivakara from Sripati. Nothing other than the *cakravāla* verses has survived of Jayadeva's mathematics and he himself is just a name, mentioned once and never again. Remembering that the depradations of Mahmud of Ghazni in the first quarter of the 11th century had left the western part of north India in chaos, it is not unreasonable to speculate that he, as well as Sripati and Udayadivakara, lived either in eastern India or to the south of the Vindhya range (like Bhaskara II a hundred years later).

In terms of motivation, *cakravāla* is a logical extension of Sripati's (and Brahmagupta's) attempts to find integral solutions for $C = 1$: begin with an integral solution for $C \neq 1$ and 'reduce' it. For $C = -1, \pm 2, \pm 4$ it could be done by *bhāvanā* alone. For general C, the technique devised by Jayadeva can be broken down into the following steps.

1. Assume that an integral solution (x_0, y_0) for some interpolator C has been found. In practice (going by the numerical examples), this was done, for a given N, by deciding on a pair (x_0, y_0) and picking $C = C_0$ to be $y_0^2 - Nx_0^2$. Compose the triple $(x_0, y_0; C_0)$ with the trivial triple $(1, v; v^2 - N)$, v an integer, coming from the solution of the trivial equation $N + v^2 - N = v^2$. The result is that $(vx_0 + y_0, Nx_0 + vy_0)$ solves the equation with $C_0(v^2 - N)$ as the interpolator:

$$N(vx_0 + y_0)^2 + C_0(v^2 - N) = (Nx_0 + vy_0)^2$$

or

$$N\left(\frac{vx_0 + y_0}{C_0}\right)^2 + \frac{v^2 - N}{C_0} = \left(\frac{Nx_0 + vy_0}{C_0}\right)^2,$$

the last being a step familiar from the study of rational solutions. It is striking that this particular *bhāvanā* of which one 'factor' is the trivial triple corresponding to a variable second root v seems never to have been employed earlier. (The one other instance of its use is later, by Bhaskara II in his solution of $ax + c = y^2, bx + d = z^2$).

2. Choose v such that

$$x_1 = \frac{vx_0 + y_0}{C_0}$$

is integral. That can always be done, by pulverising the linear Diophantine equation $x_0 v + y_0 = C_0 x_1$ with v and x_1 as the unknowns, and there is an infinite set of such pairs (v, x_1). This is obviously the breakthrough step; it bypasses the rational solution route altogether and is made possible by the special *bhāvanā* of step 1.

3. Define

$$y_1 = \frac{Nx_0 + vy_0}{C_0}$$

and

$$C_1 = \frac{v^2 - N}{C_0}$$

so that we have a new *vargaprakṛti* equation (the reduced equation)

$$Nx_1^2 + C_1 = y_1^2.$$

The bonus now is that y_1 (and hence C_1) can be made integral without losing generality; more precisely, the statement is that y_1 is integral for integral x_1 if x_0 and y_0 are mutually prime. The reason why the last condition is not a drawback is that if D is a common divisor of x_0 and y_0, then D^2 must divide C_0; we can then replace the original equation by one in which the interpolator is C_0/D^2 without disturbing the integrality of the solution. The texts do not mention this restriction. They do not also tell us how to establish the fact that an integral solution of the original equation leads necessarily to an integral solution of the reduced equation. Bhaskara II, for instance, says simply that the reduced equation (and repeated reductions) will have "two integral (*abhinna*) roots". Narayana more or less repeats the same words. (See further down for a possible justification of this neglect).

4. In the absence of guidance in the original texts, modern historians from the 19th century onwards have had to supply their own proofs of integrality. Here is one which is very natural and in the spirit of the algebraic culture of Brahmagupta and his arithmetical descendants. From the expressions for x_1 and y_1, it is easily seen that the following relations hold:

$$y_0 x_1 - x_0 y_1 = \frac{y_0^2 - Nx_0^2}{C_0} = 1,$$

$$y_0 y_1 - Nx_0 x_1 = v\frac{y_0^2 - Nx_0^2}{C_0} = v.$$

Since i) v is integral by choice, ii) x_0 and y_0 are integral by assumption and iii) x_1 is integral from the *kuṭṭaka*, it follows that x_0y_1 and y_0y_1 are both integers. If y_1 is fractional, its denominator must therefore divide both x_0 and y_0 and that is not possible when x_0 and y_0 are coprime. (The faint tinge of *reductio ad absurdum* in the last line reflects the present-day style of writing; it would not have found acceptance among Indian logicians and can easily be avoided if desired). It also follows from the first of these equations that x_1 and y_1 are in turn coprime.

5. Step 3 is iterated, giving rise to a sequence of equations $Nx_n^2 + C_n = y_n^2$ where x_n is an integer solving the linear equation $x_{n-1}v + y_{n-1} = C_{n-1}x_n$ and y_n and C_n are determined as $y_n = (Nx_{n-1} + vy_{n-1})/C_{n-1}$; $C_n = (v^2 - N)/C_{n-1}$. In principle, v can be varied from step to step. Since the goal of the iteration is to reduce $|C_n|$ ultimately to 1, it would seem advisable to choose v so as to make $|v^2 - N|$ as small as possible. Bhaskara only says that $|v^2 - N|$ should be taken small (*alpa*); it need not be the case that $|C_n| < |C_{n-1}|$ at each stage.

6. It is asserted now that the iteration will eventually produce an equation with interpolator $\pm 1, \pm 2$ or ± 4, with the prescription that if it is not unity *bhāvanā* should be employed (as described earlier). The claim thus is that, in all cases, the cyclic algorithm when combined judiciously with the principle of composition will produce a solution of the special equation. Neither Bhaskara II nor Narayana has a proof of the claim nor an acknowledgement of the need for one. The claim is true: *cakravāla* always terminates at unit interpolator. But it is doubtful whether a satisfactory demonstration was within their grasp.

In practical use, the starting interpolator C_0 of the trial equation $Nx^2 + C_0 = y^2$ is generally chosen with a trial first root $x_0 = 1$ in mind and then fixing C_0 as the smallest integer (in absolute value) that will make $N + C_0$ a square, i.e., with $(1, \sqrt{N + C_0})$ as the trial solution. Among other conveniences, we do not then have to worry about the coprimality condition, which is perhaps the reason why it is not emphasised in the texts. Bhaskara II follows the enunciation of the algorithm with two examples, $67x^2 + 1 = y^2$ and $61x^2 + 1 = y^2$. The corresponding trial equations then have interpolators $C_0 = -3$ and 3 respectively with $(x_0, y_0) = (1, 8)$ in both examples. The second of these became famous in Europe later as one of Fermat's challenge problems. Its *cakravāla* solution proceeds, briefly, as follows. By step 2, v satisfies the equation $61((v + 8)/3)^2 + (v^2 - 8)/3 = ((1 + 8v)/3)^2$. To make v integral, solve $v + 8 = 3x_1$ for integral x_1. From *kuṭṭaka*, we know (see Chapter 7.2) that its general solution is $v = 3k + 1$ with $v = 7$ as its value minimising $|v^2 - 61|$. Correspondingly, $x_1 = (7 + 8)/3 = 5$, $y_1 = (61 + 7 \times 8)/3 = 39$ and $C_1 = (49 - 61)/3 = -4$. At this point we can apply Brahmagupta's formulae for interpolator -4 given earlier in this section (involving polynomials of degree 6 in x_1 and y_1) and carry out the computation to find that $x = 226{,}153{,}980$ and $y = 1{,}766{,}319{,}049$ solve $61x^2 + 1 = y^2$. We could have, alternatively, carried on with the *cakravāla* but the size of the final numbers suggests that it will be a long calculation. It is also amusing to try other values of v: for $v = 1$, $C_1 = 20$

(with $x_1 = 3$ and $y_1 = 23$) and for the next value $v = 10$, $C_1 = 13$ (with $x_1 = 6$ and $y_1 = 47$).

In these examples (and in those from Narayana) the role of *kuṭṭaka* as a theoretical tool is minimal; the coefficients involved are small and the solution for v can be easily written down by inspection. Nevertheless, Jayadeva is quite explicit that it is by the use of *kuṭṭaka* that one finds the appropriate value of v. The term *kuṭṭaka* or *kuṭṭākāra* occurs freely in Udayadivakara's quotation from him, in contrast to Bhaskara and Narayana who do not mention the word in their exposition of *cakravāla*, maybe because it was understood by all that that was the way to go or maybe because it was trivial to find the optimal v empirically. In any case, Jayadeva's verses, even more than Bhaskara's, are clear evidence that what we consider today to be the key step in the general solution of *vargaprakṛti*, the bringing together of the qudratic and linear Diophantine equations, was also the key step for him. The inference that the interest was primarily theoretical and abstract rather than any computational need seems inescapable.

Methodologically, equally difficult to miss is the strong recursive influence in the design of the algorithm. This is not the same as saying that the method as a whole, or any of its component elements, is an example of mathematical induction in the strict sense. The recursion stops after a finite number of stages, at which point the problem is solved; in logical terms, the validity of the result does not depend on the induction axiom. As in *kuṭṭaka*, the approach is based on a reduction to a simpler case and resembles what later, in the number theory of Fermat, came to be called the method of descent. Its mathematical roots – one can invoke precedents from grammar – are to be found in Pingala, specifically in his answers to the questions 2 and 4 (Chapter 5.5), and, in a truly fundamental sense, in the reduction of a number to its decimal form by means of the division algorithm.

A final historical remark is about the several missing pieces of proof that we would have liked to see in the texts, for instance the justification for the assertion that *cakravāla* always terminates at $|C| = 1, 2$ or 4 or, more sharply and even better, at $C = 1$. A still more ambitious query: having known of the existence of an infinite number of solutions, did Jayadeva and Bhaskara II give thought to the question of whether there was a particular way of constructing them that exhausted all solutions? Such issues were satisfactorily settled in Europe only in the 18th century, when Lagrange's famous 'Addition' to Euler's work on 'Pell's' equation gave the subject a systematic theoretical framework going well beyond the clever but *ad hoc* methods of solution of the 17th century leading, among other results, to a characterisation of the minimal solution (X, Y) in terms of continued fraction expansions of $\sqrt{N}$ and the theorem that all solutions (X_n, Y_n) are generated by (X, Y) through the formula $(\sqrt{N}X+Y)^n = \sqrt{N}X_n + Y_n$. And, in the 19th centuriy, the topic got absorbed into the theory of binary quadratic forms as part of the resurgence of formal algebra and its use in number theory. In this light, what is truly striking is that Brahmagupta and Jayadeva went as far as they did by the imaginative exploitation of no more

than the basic algebraic identities satisfied by polynomials of low degree known, primarily in their geometric manifestation, from the time of the *Śulbasūtra*. The interplay between algebraic identities and Diophantine mathematics is going through a vigorous and exciting rejuvenation even as this is being written.

8.4 A Different Circle Geometry: Cyclic Quadrilaterals

At first sight, the other field in which Brahmagupta made singularly original contributions, namely the geometry of cyclic quadrilaterals, may seem to have not much to do with his number theory. The geometry comes before the arithmetic and algebra in the *Brāhmasphuṭasiddhānta*, as part of Chapter XII, and it comes in two distinct flavours: one, numerical geometry, which culminates in the construction of cyclic quadrilaterals whose sides and other associated magnitudes are rational or integral numbers, and the other concerned with general ('real') cyclic quadrilaterals. It becomes quickly evident however that the integral geometry is actually the geometric expression of his preoccupation with what may be called quadratic number theory; several of the propositions of Chapter XII attest to this linkage. It is also possible to make a case that his results on general cyclic quadrilaterals drew their inspiration from his very detailed understanding of their integral counterparts or, at least, that the two flavours of geometry were not isolated one from the other in his mind, rather as the *Śulbasūtra* did not separate into different compartments the geometric theorem of the diagonal from the geometric consequences of the existence of rational diagonal triples. A substantial portion of the geometry section of Chapter XII is in fact devoted to rational triangles and figures that can be built out of them.

The first striking fact about Chapter XII is that it does not have any Aryabhatan trigonometry. That is the reason for the title of this section; it really is a direct outgrowth and a renewal of the *Śulbasūtra* view of geometry. Trigonometric functions are banished to the astronomical chapters where they are put to immediate use; it is as though the new-fangled stuff about half-chords and their values was not as authentically geometric as the knowledge that could be traced back to Vedic times. Surprisingly, Bhaskara II did the same five centuries later; numerical geometry of cyclic figures, including the computation of the full chords making regular polygons of up to 9 sides, is in *Līlāvatī* while the sine table and everything connected with it is in the *Golādhyāya*, the chapter on the sphere in the *Siddhāntaśiromaṇi*. The reasonable conclusion is that cyclic polygons and their properties were thought to have little connection with trigonometric functions, a misperception that was finally set right by Madhava and his school.

Bhaskara II was clearly much taken with Brahmagupta's numerical geometry, generalising and illustrating profusely several of the latter's constructions. But it is also well documented that he was not aware of Brahmagupta's

unstated restriction (though it is implicit in the terminology, see later in this section) that the quadrilaterals for which his theorems on diagonals and areas held were necessarily cyclic. After expressing his mystification at the area of a quadrilateral being fully determined by its sides – since the diagonals of a non-cyclic quadrilateral are not so determined – there is his famous tirade against persons propagating such errors (and those who fall for them) as contemptible demons. Unstated assumptions and unexplained terms have always been part of the written mathematical corpus in India – the price paid for conciseness – and such incompleteness had never been damaging to faithful communication. Perhaps there was a break in the transmission of tradition, understandable if it was dependent on face-to-face instruction. Perhaps, as has been suggested, Bhaskara's invective was aimed not at Brahmagupta but at others, ignoramuses who did not understand the distinction. Anyhow, the true inheritor of Brahmagupta's real circle geometry was not Bhaskara but Narayana who introduced a genuine conceptual novelty (that of the 'third diagonal') arising from the symmetries of a general cyclic quadrilateral, leading to beautiful new theorems about them. For that reason, an account of Narayana's geometric work belongs in this chapter (the next section) though he is closer in time to the Nila school than to Bhaskara. Indeed, Narayana's results led on to some of the trigonometric reinterpretations of the new circle geometry due to the Nila mathematicians; we will see examples in Part III.

Much of Brahmagupta's rational and integral geometry consists of fairly straightforward applications of the diagonal theorem and some of it could well have been accomplished by the *Śulbasūtra* geometers; conceptually and in some of the specific results, the connection is very close indeed. Rational scaling of right triangles, first utilised by Apastamba and stated as a general property by Manava in their versions of the *Śulbasūtra* (see Chapter 2.2), is put to very productive use in many of the constructions. Moreover, the passage from rational to integral geometry is easy: scale the appropriate right triangles suitably and then clear denominators. Facilitating such constructions is the general formula for diagonal triples as generated by a pair of integers; in the *Brāhmasphuṭasiddhānta* it takes the form of the geometric statement that, in an (integral) isosceles triangle, the symmetric sides are $k^2 + l^2$, the altitude is $2kl$ and the base is $2|k^2 - l^2|$, for any two numbers $k \neq l$ (Brahmagupta does not mention the other, geometrically distinct, assignment interchanging the altitude and the half-base); we recognise it as just the generalisation of Katyayana's formulae for the same quantities, in his procedure for multiplying a square by n (Chaper 2.3). And underlying it all is the same quadratic identity, $(x + y)^2 = (x - y)^2 + 4xy$, that plays such an important, though sometimes hidden, role in the quadratic Diophantine equation.

Because Brahmagupta's prescription for integral cyclic quadrilaterals of a special type is the capstone of this line of thinking, it is the only integral-geometric topic I discuss; and because it may have been the bridge that led to the study of real cyclic quadrilaterals in general, I take it up first (inverting the order in the text).

The term ‘Brahmagupta quadrilateral’ as commonly used today often includes in its definition several properties which Brahmagupta clearly thought of as consequences of his basic construction. Here is the relevant verse:

> The uprights and the bases (i.e., the two perpendicular sides) of two generated (*jātya*) [right triangles] are to be multiplied by each other’s hypotenuse. They are the four sides of a scalene (*viṣama*, no two sides equal) [quadrilateral]. The longest is to be the base [of the quadrilateral], the shortest the face (the side opposite the base) and the other two the flanks.

As can be seen from the square brackets, there are the usual terminological short cuts. The verse does not even say that the four sides are integral; that is to be deduced from the context and the use of the word *jātya* to mean an integral right triangle, possibly referring to its being generated by (literally, ‘born out of’) the pair (k, l) as the diagonal triple $(k^2 + l^2; k^2 - l^2, 2kl)$. It is not explicit (and not immediately evident, for the ordering of the sides as prescribed) that the quadrilateral is cyclic. These apparent ambiguities get cleared up once we look for the underlying motivation for the cross multiplication of the sides of the *jātya triangles* by each other’s hypotenuse. This was brought out clearly by Bhaskara II (he had no doubt that, in this instance, the quadrilateral in question was indeed cyclic) who then derived several corollaries following from Brahmagupta’s construction. The steps involved are as follows.

The key idea is the old one of the scaling of right triangles. The purpose of the reciprocal multiplication can only be to scale the two triangles so as to make their hypotenuses equal. They can then be brought together with the hypotenuses coinciding, without any overlap in the areas. The resulting quadrilateral is obviously cyclic since one pair of opposite angles is each $\pi/2$. If $(h; p, q)_i$, $i = 1, 2$, are the triples characterising the two triangles, the sides of the quadrilateral are p_1h_2, q_1h_2, p_2h_1 and q_2h_1 and the diameter of the circumcircle, which is also the length of one diagonal, is $\mathcal{D} = h_1h_2$ (see Figure 8.1). Flipping the alignment of one hypotenuse will switch the pair of sides

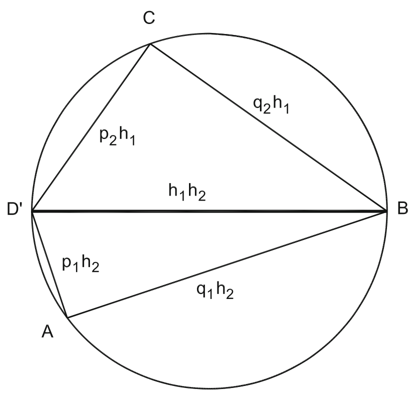

Figure 8.1: The pre-Brahmagupta quadrilateral

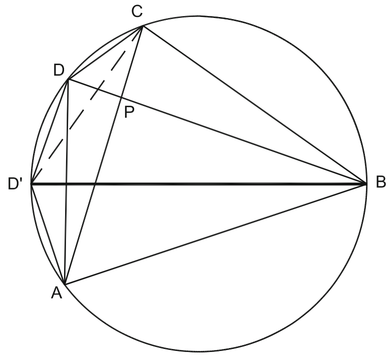

Figure 8.2: The Brahmagupta quadrilateral and its diagonals

in one semicircle. The longest and shortest sides of the quadrilateral will be adjacent and perpendicular, and their squares will add up to $\mathcal{D}^2$; the same is true of the two other sides. So this is not yet the quadrilateral described by Brahmagupta. For ease of reference, it is useful to give it a name and I will call it a pre-Brahmagupta quadrilateral. To bring the longest and the shortest sides opposite one another, it is sufficient now to shift the vertex D' to D such that $CD = AD'$ (and $AD = CD'$), i.e., to interchange or transpose the chords AD' and CD' keeping A and C fixed. $ABCD$ is then a Brahmagupta quadrilateral according to his instruction. The operation of transposition does not change the perimeter or the area. The following properties are immediate: i) the squares of each pair of opposite sides sum up to $\mathcal{D}^2$; ii) area $(ABCD)$ $(=$ area $(ABCD'))$ $=$ half the sum of the products of opposite sides; and iii) the chord $D'D$ is parallel to the (unchanged) diagonal AC. The last is a crucial property of transpositions in general, made much use of in later work. Here it implies, from the orthogonality of $D'D$ and BD, that the diagonals AC and BD intersect orthogonally (Figure 8.2). Moreover, the lengths of the diagonals are computable in terms of the sides $(p, q)_i$ of the two *jātya* triangles. The geometric way to do this is to note that the right triangles APB and DPC are similar:

$$\frac{AP}{DP} = \frac{BP}{CP} = \frac{AB}{CD} = \frac{q_1 h_2}{p_1 h_2} = \frac{q_1}{p_1}$$

and that the right triangles APD and BPC are similar:

$$\frac{AP}{BP} = \frac{DP}{CP} = \frac{AD}{BC} = \frac{p_2}{q_2}.$$

The lengths of the diagonal segments are however related, $AP^2 + BP^2 = AB^2 = (q_1 h_2)^2$ and so on, and that fixes each of them: $AP = q_1 p_2$, $BP = q_1 q_2$, $CP = p_1 q_2$ and $DP = p_1 p_2$. The full diagonals are therefore

$$d_1 = AC = p_1 q_2 + q_1 p_2 \qquad d_2 = BD = p_1 p_2 + q_1 q_2$$

in a form first written down by Bhaskara II. To these we may add two results already noted: the circumdiameter is

$$\mathcal{D} = h_1 h_2 = \sqrt{(p_1^2 + q_1^2)(p_2^2 + q_2^2)}$$

and the area is

$$\mathcal{A} = \frac{1}{2}(p_1 h_2)(q_1 h_2) + \frac{1}{2}(p_2 h_1)(q_2 h_1) = \frac{1}{2}(p_1 p_2 + q_1 q_2)(p_1 q_2 + q_1 p_2)$$

which is $(1/2)d_1 d_2$.

The Brahmagupta quadrilateral is thus characterised by its geometric property of orthogonal diagonals as well as the arithmetical relationships, arising from the geometry, among various natural lengths and areas, all integral (of which the quantities I have computed above are a sample). My description is closest to that of Bhaskara II (in *Līlāvatī*). In particular, he considers all quadrilaterals obtained from a pre-Brahmagupta quadrilateral by repeated transpositions of pairs of sides (i.e., by general permutations of sides) and states that they fall into two classes (easy to show): either one diagonal is a diameter (a pre-Brahmagupta quadrilateral as in our initial step) or the two diagonals are perpendicular (the unique, up to orientation, Brahmagupta quadrilateral). As the cited verse from the *Brāhmasphuṭasiddhānta* shows, Bhaskara's interpretation, in particular the operative role of transpositions of sides, can reasonably be taken as reflecting Brahmagupta's own thinking. Bhaskara emphasised the geometry of *jātya* triangles but there was a lot of work in the subject before him, most comprehensively by Mahavira, that relied on fairly involved algebraic manipulation of the pairs $(k, l)_i$ generating the triangles. (Bhaskara has critical things to say about his predecessors' complicated method of computing the diagonals). It is possible that their reliance on algebraic identities as a tool of geometry is also a reflection and a continuation of Brahmagupta's own predilection for numbers above all else. (It will become apparent presently that this statement is not in contradiction with his theorems on real cyclic quadrilaterals). Such a bias may also account for his exclusion of non-cyclic quadrilaterals from consideration. The latter are not easily amenable to an arithmetical approach, the link between geometry and arithmetic provided by the diagonal theorem and diagonal triples, so productively put to use *via* the *jātya* triangles, being no longer immediately available.

A minor historical remark is that the intial step of juxtaposing two right triangles to make a cyclic quadrilateral depends on the fact that the diameter subtends a right angle at any point on the circle. The proposition, sometimes called Thales' theorem, is in Euclid (Book III in which the geometry of the circle is initiated) but not found nor used in the *Śulbasūtra* or the *Āryabhaṭīya*. But this 'Greek' orthogonality – as opposed to the 'Indian' orthogonality of intersecting diameters and half-chords – is only a small step from Aryabhata's proposition on segments of the diameter (*Gaṇita* 18, see Chapter 7.1): in Figure 8.3, $AC^2 + BC^2 = AP^2 + BP^2 + 2CP^2 = AP^2 + BP^2 + 2AP \cdot BP = AB^2$.

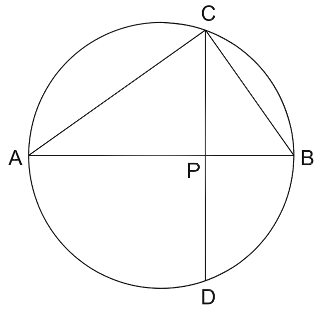

Figure 8.3: 'Greek' and 'Indian' orthogonality

It would be rash to suppose that Brahmagupta (or others before him) could not have taken that step; in fact, there is an astronomical application of the theorem in his late work *Khaṇḍakhādyaka*. Subsequently both Narayana and the Nila mathematicians made good use of Thales' theorem in their extensive investigations of the general cyclic quadrilateral.

None of the various interrelationships found above depends on the integrality of Brahmagupta quadrilaterals. That is obvious geometrically and also in the algebraic approach since the identities employed, which are reformulations of the diagonal theorem, are all valid for real numbers. This trivial observation lets us speak of a real (the qualification can be dropped from now on) Brahmagupta quadrilateral as a cyclic quadrilateral satisfying the additional condition that its diagonals intersect orthogonally. (Any one of the several equivalent geometric properties that hold in the integral case can be imposed instead; orthogonality of diagonals is the most efficient). The sides, $a_1 = AB$, $a_2 = BC$, $a_3 = CD$, $a_4 = DA$, are now real numbers (with a_1 as the longest side and a_3 as the shortest) as are the diagonals $d_1 = AC$ and $d_2 = BD$, the circumdiameter $\mathcal{D}$ and the area $\mathcal{A}$.

The point of freeing the Brahmagupta quadrilateral from the integrality condition has to do with his two celebrated theorems on general (real, not necessarily Brahmagupta) cyclic quadrilaterals. In the order in which they occur in the *Brāhmasphuṭasiddhānta*, they are:

1. The area theorem: If $2s = a_1 + a_2 + a_3 + a_4$ is the perimeter, then the area is given by
$$\mathcal{A}^2 = (s - a_1)(s - a_2)(s - a_3)(s - a_4).$$

2. The theorem of the two diagonals or, simply, the diagonal theorem (no confusion with the Vedic diagonal (Pythagoras') theorem is likely to arise): The lengths of the two diagonals are given by
$$d_1^2 = \frac{a_1a_4 + a_2a_3}{a_1a_2 + a_3a_4}(a_1a_3 + a_2a_4),$$
$$d_2^2 = \frac{a_1a_2 + a_3a_4}{a_1a_4 + a_2a_3}(a_1a_3 + a_2a_4).$$

As already noted, a special case of the area theorem is Heron's formula for the area of the triangle, $\mathcal{A}^2(a_4 = 0) = (s - a_1)(s - a_2)(s - a_3)s$ and the diagonal theorem has Ptolemy's theorem, $d_1 d_2 = a_1 a_3 + a_2 a_4$, as a corollary.

The formula for the area is in fact the first geometric proposition in Chapter XII. Apart from its mathematical content, the verse is an instructive example of the problems that beset the *sūtra* style of writing and of how they may (though not always) be resolved. The geometric object named in it is *tricaturbhuja* and most modern commentators starting with the first translator, Colebrooke ([Co]), have adopted the analysis of the compound word as *tribhuja* and *caturbhuja*, triangles and quadrilaterals, a reading that goes all the way back to Brahmagupta's distinguished commentator Prthudaka (10th century). That interpretation is incorrect from the context alone, never mind the grammar. (When Aryabhata writes about both triangles and quadrilaterals at the same time, he is very explicit: *tribhujā-ccaturbhujā* (*Gaṇita* 11) which is the canonical conjunctive composition). The verse speaks of a side *bāhu* and the opposite side *pratibāhu* of this object and that of course makes no sense for a triangle. And, in the statement of the formula itself, there is no ambiguity: we are to take the semiperimeter and subtract from it each of the four sides, multiply the results and then take the square root; if there were only three sides to be subtracted, we would presumably have been told what to do with the fourth factor. It is only very recently that the suggestion has been made, supported by a suitable deconstruction of the compound, that the word stands for a quadrilateral that shares a property of all triangles, that of being circumscribable by a circle.[6] An even simpler reading of *tricaturbhuja* is to take its meaning, a 'three-sided quadrilateral', literally; it is grammatically natural and mathematically appropriate: a cyclic quadrilateral has only three independent sides, the fourth being determined by them. (We will see in the following that the three-sidedness is a running theme in the Indian approach to cyclic quadrilaterals, from Brahmagupta to *Yuktibhāṣā*). In any case, independently of any grammatical analysis, there is no doubt possible that Brahmagupta meant these two theorems, as well as all his other results on *viṣama* quadrilaterals, to be valid only for cyclic quadrilaterals (and for all of them). Narayana and the Nila mathematicians had no confusion at all about what *tricaturbhuja* meant.

These results aroused somewhat less interest than Brahmagupta's integral geometry in the centuries that followed, it is not obvious why, until we come to Narayana. In fact complete proofs, building on Narayana's extensive investigations of general cyclic quadrilaterals, were written down even later, in *Yuktibhāṣā*. The proofs are quite elaborate; the steps involved appear at first sight to be a matter of ingenuity and imagination, with no obvious hint that allows an assessment of what they might have owed to Brahmagupta's own path

[6]Pierre-Sylvain Filliozat, "Modes of Creation of a Technical Vocabulary: the Case of Sanskrit Mathematics", Gaṇita Bhāratī **32** (2010) p. 37. The analysis is not very direct. Having invented a new term and exemplified its meaning at the beginning, Brahmagupta hardly uses it again (there is one other verse where its meaning is just as unambiguous) contenting himself most of the time with the adjectival substitute *viṣama* (scalene).

to the theorems. For we cannot doubt that he must have had proofs for them or, at the very least, sound reasons for believing in their truth – they are sufficiently non-intuitive and sufficiently in advance of contemporary mathematical knowledge to exclude the possibility of mere inductive guesswork. (We shall come to the question of possible external influences presently). What might the reasoning have been? In the absence of definitive documentary evidence, it seems worth asking whether his work on the special case of Brahmagupta quadrilaterals might not have acted as a guidepost to the general case. As it turns out, the formulae for the diagonals are surprisingly easy to derive in the special case and the area formula slightly less direct but not much harder. More importantly for the history of the subject, we can then see how Narayana's whole approach – and, hence, the Nila proofs – is an organic development of techniques that work so effortlessly when the diagonals intersect orthogonally. Later on I will mention another recent reconstruction of how Brahmagupta's own proof might have gone. We will perhaps never know for sure his original line of reasoning; whatever it might have beeen, the fact that these beautiful theorems are not difficult to establish for Brahmagupta quadrilaterals seems not to have been put on record before. The broadening and deepening of an initial discovery so as to fit it into a more general setting is a common enough phenomenon in mathematics: two examples that readily come to mind, from either end of the chronological span of Indian mathematics, are the generalisation of the diagonal theorem from the square to the rectangle and the passage from Madhava's π series to the general arctangent series.

The expressions for the diagonals of a Brahmagupta quadrilateral in terms of the sides $(p,q)_i$ of the *jātya* triangles (Bhaskara II's formulae) are easily rewritten as functions of the sides of the quadrilateral ($a_1 = p_1 h_2$, etc.): $d_1 = (a_1 a_4 + a_2 a_3)/(h_1 h_2)$, $d_2 = (a_1 a_2 + a_3 a_4)/(h_1 h_2)$. Their ratio is

$$\frac{d_1}{d_2} = \frac{a_1 a_4 + a_2 a_3}{a_1 a_2 + a_3 a_4}.$$

But we also know their product

$$d_1 d_2 = 2\mathcal{A} = a_1 a_3 + a_2 a_4,$$

the last equality resulting from the fact that (a_1, a_3) and (a_2, a_4) are the sides of the scaled *jātya* triangles from which the corresponding pre-Brahmagupta quadrilateral was built up and from the invariance of the area under transposition of sides – showing, in passing, that Ptolemy's theorem is trivial for a Brahmagupta quadrilateral. From their ratio and their product, the formulae for the individual diagonals follow and they are the same as Brahmagupta's formulae for the general cyclic quadrilateral; orthogonality of the diagonals introduces no extra simplification. For a proof of the general theorem, we have therefore only to show how and why these formulae survive the generalisation.

To be noted is that d_1 is unchanged under transpositions of a_1 with a_4 and a_2 with a_3, and d_2 correspondingly. These are symmetries which hold for the diagonals of the general cyclic quadrilateral as well.

For the area, we already have an expression $\mathcal{A} = (a_1a_3 + a_2a_4)/2$ in terms of the sides. But it cannot cover the general case as it stands since it is not invariant under transpositions of sides. What we can try to do to is to restore the invariance, i.e., symmetrise the expression for the area, by exploiting relationships that hold between the sides of the Brahmagupta quadrilateral. The only such is the equality of the sums of squares of opposite sides, $a_1^2 + a_3^2 = a_2^2 + a_4^2$ $(= \mathcal{D}^2)$, which, to remind ourselves, comes from the construction of the pre-Brahmagupta quadrilateral. To make use of it, consider the identity

$$(a_1 + a_3)^2 - (a_2 - a_4)^2 = a_1^2 + a_3^2 - a_2^2 - a_4^2 + 2(a_1a_3 + a_2a_4).$$

The left side of this equation is identically $(a_1+a_3+a_2-a_4)(a_1+a_3-a_2+a_4) = 4(s-a_2)(s-a_4)$. For the Brahmagupta quadrilateral, the four square terms on the right side add up to 0 and what is left is $4\mathcal{A}$. But we also have, in the same way, $4(s - a_1)(s - a_3) = (a_2 + a_4)^2 - (a_1 - a_3)^2 = a_2^2 + a_4^2 - a_1^2 - a_3^2 + 2(a_1a_3 + a_2a_4) = 4\mathcal{A}$. On multiplying the two expressions for $\mathcal{A}$, we get the area of the Brahmagupta quadrilateral in a form that respects transposition invariance and it is the same as his formula for the general cyclic quadrilateral. To establish it in the general case, it is therefore enough to show that the product of the right sides of these two identities is still $(4\mathcal{A})^2$ even though the first four terms will not sum up to 0:

$$(4\mathcal{A})^2 = [(a_1^2+a_3^2-a_2^2-a_4^2)+2(a_1a_3+a_2a_4)][(a_2^2+a_4^2-a_1^2-a_3^2)+2(a_1a_3+a_2a_4)].$$

This (or some rearrangement of it) is exactly what the later proofs establish.

The purpose of this little exercise was to put forward the case that Brahmagupta had, at the very least, strong grounds for believing in the validity of his general theorems – that they were good conjectures in modern parlance – and that Brahmagupta quadrilaterals provided that ground. The algebraic identities employed to symmetrise the area bear his unmistakable signature – compare the additive and subtractive identities that are the key to his *bhāvanā* – as does the idea of the invariance of the area under transpositions. There have been suggestions in the past that he did not consider the general cyclic quadrilateral at all but meant his theorems to hold only for those with orthogonal diagonals. If that indeed were the case, it would be difficult to understand why he bothered with the more general symmetric expression for the area when there was available the simpler (exact) formula $(a_1a_3 + a_2a_4)/2$. Indian commentators never had that doubt. The central idea in going from the special to the general case is that the invariance of the area under transpositions holds for the general cyclic quadrilateral; having noted and exploited it in the making of his eponymous quadrilateral, Brahmagupta could not possibly have missed its general validity. As for the pervasiveness of algebraic identities, and algebraic reasoning more generally, that has always been a characteristic of Indian geometry from the time of the *Śulbasūtra*; the general proofs of the two theorems, when they were eventually set down, were just as heavily algebraic and just as reliant on the key idea of transposition invariance, particularly in the form Narayana gave to it.

Before moving on to Narayana's approach to cyclic quadrilaterals (and what it owed to earlier ideas regarding the role of transpositions) and to the proofs (in *Yuktibhāṣā*), I deal briefly with some other points, connected one way or another with tracing Brahmagupta's route to his theorems and his putative proofs for them.

The first is primarily of historical interest. Almost alone among Indian geometrical results, each of Brahmagupta's two theorems had clear precedents in the form of a partial statement or a special case, both from Alexandria. The formula for the product of diagonals

$$d_1 d_2 = a_1 a_3 + a_2 a_4$$

is Ptolemy's theorem, the main tool in the compilation of his chord table. And the formula for the area of the quadrilateral specialises as

$$\mathcal{A}^2(a_4 = 0) = (s - a_1)(s - a_2)(s - a_3)s$$

for the area of the triangle, attributed to Heron. Chronologically and on the face of it, the case for a causal connection is far stronger than it is for Aryabhatan trigonometry (see Chapter 7.3). Nevertheless, the question of whether Brahmagupta was aware of these Alexandrian partial anticipations is not easy to answer categorically. Though intellectual contacts between Hellenic Egypt and northwest India ceased after the destruction of Taxila in mid-5th century CE, knowledge exchanged prior to that event may well have survived, as in so much of the astronomy. But, as we have seen earlier, and in contrast to astronomy, there is nothing in Indian geometry that we can identify unambiguously as of Greek origin (excluding of course the common ground that all geometrically aware ancient cultures shared). The triangle itself has a generally subsidiary role in India except when it is the diagonally divided half of the rectangle. The area of the general triangle is always given, at least from the time of Aryabhata, as the product of half the base and the altitude or variations or specialisations thereof. Heron's formula occurs nowhere prior to the writing of *Yuktibhāṣā* though that meaning has sometimes been read into Brahmagupta or Bhaskara II by modern writers who, mistakenly, take *tricaturbhuja* to stand for both a triangle and a quadrilateral. In fact the *Yuktibhāṣā* proof of Heron's formula is not as a trivial special case of Brahmagupta's quadrilateral formula but as an independent but parallel derivation (and it is very different from the Greek proof). All in all, it does not seem as though the triangle formula was a 'known' which was then generalised to the quadrilateral. We have to conclude that, though the chronolgy is not against transmission, mathematical evidence for a causal relationship is absent.

The situation regarding Ptolemy's theorem is not very different. It is an obvious property of Brahmagupta quadrilaterals (see above), but it is not mentioned or used by anyone (with the possible exception of the commentator Prthudaka), either as a corollary of Brahmagupta's diagonal theorem or independently of it, before (again, possibly) Narayana. But the fact is that if its

knowledge is assumed, the diagonal theorem can be given an extremely simple proof. As we shall see in the next section, it is elementary to prove that the formula for the ratio of the diagonals of the Brahmagupta quadrilateral in terms of its sides (see above) remains valid in the general case. Ptolemy's theorem says that that is also true of their product (though it is no longer twice the area), thus determining the individual diagonals. In any case, whether there was an Alexandrian input or not, the Indian approach to diagonals is distinctively structural-algebraic, very different from and more elaborate than the Greek proof(s) of Ptolemy's theorem.

One might even say that the interest in areas and diagonal lengths of quadrilaterals more complicated than the rectangles of the *Śulbasūtra* is the most fundamental manifestation of the indigenous inspiration of Brahmagupta's geometry, specifically his affinity with Vedic geometry: the square and the rectangle after all are the simplest pre-Brahmagupta quadrilaterals. There are other tell-tale signs. Brahmagupta's word for the radius is the archaic-sounding *hṛdaya-rajju*, the cord to the heart or the centre (not the *viṣkambhārdha*, half-diameter, of Aryabhata), strongly evocative of the *rajju* used in the *Śulbasūtra* as the radial cord for drawing a circle. And, as with the quadratic Diophantine equation, his geometrical results were of no use in astronomy and led to no trigonometrical insights in his hands (that had to wait for the Nila school); they are untouched by the Aryabhatan revolution. In that sense the new circle geometry is just the old *Śulbasūtra* geometry renewed and reinvigorated.

Indeed, there is a sequence of generalisations, from the simplest particular case to the most general, always staying with cyclic quadrilaterals: squares/rectangles (*Śulbasūtra*, one diagonal) → pre-Brahmagupta (one of two diagonals is a diameter) → Brahmagupta (two mutually orthogonal diagonals) → general (also Brahmagupta; Narayana's three diagonals).

The second point concerns a recent and extremely thorough analysis of the sequence of eight verses in the *Brāhmasphuṭasiddhānta* beginning with the area theorem and ending with the diagonal theorem.[7] The analysis concludes that these verses constitute a schema for a proof of the area theorem and, as an auxiliary result, the diagonal theorem as well. Much has been written about these geometrical verses in the past but there is little doubt now that the earlier studies were often misled by their obscurities, terminological – *tricaturbhuja* being the most critically important example – and otherwise. Kichenassamy's main point is that a careful and coherent rereading and reinterpretation of the whole passage allow the reconstruction – or, rather, the restoration – of a proof which does not call upon later texts and can therefore be wholly attributed to Brahmagupta. The case is persuasively made though one might quibble with the precise interpretation of one or two passages.

[7]Satyanad Kichenassamy, "Brahmagupta's Derivation of the Area of a Cyclic Quadrilateral", Historia Mathematica, vol. **37** (1) (2010) p.28.

Kichenassamy's view of the conceptual underpinning of Brahmagupta's approach to cyclic quadrilaterals is quite different from the perspective adopted in the present account. In particular, it has no point of contact with Brahmagupta (or pre-Brahmagupta) quadrilaterals, integral or real. Consequently, transpositions have no place in the picture that emerges (though transposition invariance of the area is used where convenient in the argumentation). Their role is given to two reflection operations applied to a scalene triangle in its circumcircle as the primary geometric object, through two mutually perpendicular lines, suitably chosen. There is not enough in the *sūtra*s for us to decide whether Brahmagupta had such reflections in mind – reflection operations, as opposed to the mere presence of bilateral symmetry, do not appear anywhere else in Indian geometry. What is certain, on the other hand, is that the construction of the Brahmagupta quadrilateral from reciprocally scaled right triangles could not have done without transposition.

Related to this is the fact that the expression for the area that Kichenassamy's derivation leads directly to is the one involving squares and products of sides that I have written down above and this is true also of the *Yuktibhāṣā* derivation. If the purpose was just to write the area as a function of the sides, Brahmagupta could have stopped there. It is then transformed through algebraic identities (also as explained above) into the elegant symmetric form involving the semiperimeter s: "Half the sum of the sides, set down four times, and severally lessened by the sides, being multiplied together, the square root of the product is the exact area" (Colebrooke's translation, [Co]). Why would he give himself the trouble of rewriting a perfectly satisfactory formula in a variant form? In order to exhibit the symmetry in as manifest a manner as possible? Or, perhaps, he actually arrived at that form, which he would have if he started with the relation $a_1^2 + a_3^2 = a_2^2 + a_4^2$ satisfied by the sides of the Brahmagupta quadrilateral and used it to symmetrise its area. We do not know.

Connected to this question is another mystery. Preceding the exact area formula (the first line of the same verse) is a rough (*sthūla*, often translated as approximate) value for it, meant, in all probability, as the starting point of his approach. The rough formula is: *bāhu-pratibāhu-yoga-dala-ghāta*, "the product of half the sums of opposite sides", $(a_1 + a_3)(a_2 + a_4)/4$. This is the correct formula only for a rectangle. Why would Brahmagupta start with such a poor first guess (*sthūla* is often used to denote the starting point in an approach to finer and finer approximations), one which was known to the authors of the *Śulbasūtra* if not earlier as exact for rectangles, and more simply expressible as just a_1a_2? If, on the other hand, we transpose *yoga* and *ghāta*, the *sthūla* formula becomes "the sum of half the products of opposite sides", $(a_1a_3+a_2a_4)/2$, exact for any Brahmagupta quadrilateral. Not only is it a better starting point, but it is also the trajectory he himself seems to have followed, if we accept *sthūla* here as meaning not any approximate value but as the initial value which is to be refined further. The implausibility of the text as it has always been read may well be the reason why no commentator pays much attention to the *sthūla* formula. All that is required to set right the anomaly is for there to

have occurred an error in the transmission of the text, which is not unheard of. Alternatively, if such an error did occur, it will strengthen the thesis proposed here of a causal connection from Brahmagupta quadrilaterals to general cyclic quadrilaterals in Brahmagupta's thinking on the subject.

What matters in the end is that we now have, thanks to Kichenassamy's work, one more good reason to believe that Brahmagupta's two theorems were meant to apply to all cyclic quadrilaterals and that they are really theorems, not mere conjectures no matter how plausible, and certainly not the result of inductive guesswork alone.

8.5 The Third Diagonal; Proofs

In regard to the geometry of cyclic quadrilaterals, Narayana's position is somewhat similar to that of Jayadeva's in regard to the quadratic Diophantine equation: he carried to a satisfactory conclusion Brahmagupta's ideas concerning transpositions of sides as symmetries that operate on them. His brilliant innovation of the third diagonal as the extra ingredient best suited for the exploration of such transformations supplied the finishing touch that turned the Indian approach to cyclic quadrilaterals into a finished whole. Propositions already known became special cases of general theorems and interesting new results emerged. Though he does not seem to have recorded them (in his known works), the proofs of Brahmagupta's two main theorems in the subject that were written down soon after his time made use of some of the concepts and techniques he introduced. As we shall see in a minute, the proof of the diagonal theorem (and of a striking generalisation of it) is only one short step from the results he stated.

Narayana's geometry is in his *Gaṇitakaumudī* (mid-14th century), a work which treats several other themes. We will return to what little we know of him, and the little more that we can guess, in Chapter 9.4, in the right chronological context. Here I confine myself just to his contributions to what we may justly call the Brahmagupta programme. The text has no proofs; those given below are based on later expositions (of which, alas, there were not to be too many – we are only about two and a half centuries from the end of mathematics in India) such as *Yuktibhāṣā*, or reconstituted in a way consistent with the extant knowledge and culture.

Once introduced, the third diagonal appears as a very natural idea, arising from the extension of transposition operations, first applied by Brahmagupta to his eponymous quadrilaterals and followed up by Bhaskara II, to general quadrilaterals ($ABCD$ in Figure 8.4); Narayana leaves no doubt that he has in mind all (*sarva*) cyclic quadrilaterals. Interchanging the sides AD and CD as before results in the transposed quadrilateral $ABCD'$ with $AD = CD'$, $AD' = CD$ and DD' parallel to AC; the auxiliary symmetric trapezium $ACDD'$ so created plays a role in the proofs of several propositions in Narayana's approach to cyclic quadrilaterals. The diagonals AC and BD are considered the first two

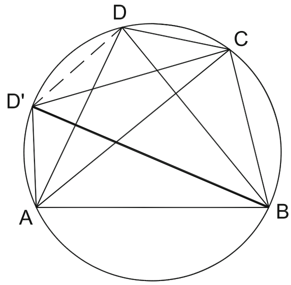

Figure 8.4: Transposition and the third diagonal

diagonals and the diagonal BD' of the transposed quadrilateral is Narayana's third diagonal. Denote their lengths by $AC = d_1$, $BD = d_2$ and $BD' = d_3$ (the sides will be denoted as earlier, $AB = a_1$, $BC = a_2$, $CD = a_3$ and $DA = a_4$).

Other transpositions (and their combinations) will result in other geometrically distinct third diagonals but their lengths can take one of only three values, d_1, d_2 or d_3. That is because, since the four sides are four chords of a fixed circle such that the corresponding arcs cover the circle without overlap, any transposition is a permutation of three of the chords and their arcs, the fourth then being fixed. (One can also draw the transposed quadrilaterals and demonstrate geometrically the uniqueness of the third diagonal, as *Yuktibhāṣā* does). We may therefore think of a given cyclic quadrilateral as having three, and only three, diagonals as far as their lengths are concerned – that is how Narayana states it – and indeed as being fully determined by them (up to ordering of sides) since there are only three independent sides. Alternatively, we may think of transpositions as generating a set of geometrically distinct quadrilaterals from a given cyclic quadrilateral, all having the same set of three distinct diagonal lengths which get permuted among themselves, an equivalence class of quadrilaterals in modern terminology. It may not appear an undue liberty to present his idea in such an abstract guise if we remember that Narayana was also a great combinatorialist, particularly strong in the use of combinations and permutations (we shall see examples later), the last in the lineage of Pingala. That he thought of interchanges as precisely defined operations is attested by the use of a specific technical term, *parivartana*, for them.

Turning to what Narayana did with his new concept, a first remark is that the third diagonal of a Brahmagupta quadrilateral is the diameter $\mathcal{D}$, as follows from the fact that it is the result of a transposition applied to a pre-Brahmagupta quadrilateral. (It is also easy to demonstrate directly that if two diagonals intersect perpendicularly, the third diagonal is a diameter). In other words, an alternative characterisation of a Brahmagupta (including pre-Brahmagupta) quadrilateral is as one which has a diameter as one of its three diagonals. Results on general cyclic quadrilaterals can therefore be specialised to Brahmagupta quadrilaterals by reexpressing them in terms of diagonals and putting one of them equal to $\mathcal{D}$. Conversely, one way to explore possible ex-

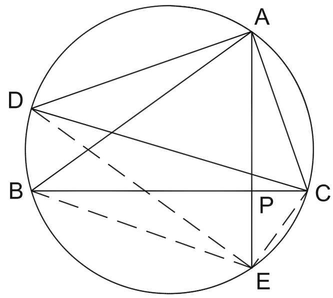

Figure 8.5: Altitude and circumdiameter

tensions of results which hold for Brahmagupta quadrilaterals to the general case is by expressing them in terms of $\mathcal{D}$ and replacing $\mathcal{D}$ by say d_3. Several of Narayana's results can be guessed at this way.

A second remark is the obvious one: the invariance of the area under transpositions is a property of all cyclic quadrilaterals. This was thought to be a sufficiently important observation as to deserve an explanation in *Yuktibhāṣā*: the area of the quadrilateral is the area of the circle less the sum of the areas bounded by each of the four sides/chords and its arc, and hence unchanged under transpositions. Much of Narayana's research on cyclic quadrilaterals is built on this invariance. There are several pretty results that came out of it, but the emphasis here will be on those that connect directly with the theorems of Brahmagupta. Other propositions not discussed here and their proofs will be found in Sarasvati Amma's book ([SA]).

An essential input in these results is an elementary relationship between an altitude of a triangle and its circumdiameter, typical of the advantages of not divorcing the triangle from its circumcircle: if AP is an altitude of the triangle ABC and $\mathcal{D}$ is its circumdiameter, then $AP = AB \cdot AC/\mathcal{D}$. Narayana does not give a proof but we can easily supply one. Let CD be the diameter through C of the circumcircle (Figure 8.5). Then ABP and CDA are similar triangles and the result follows.

It will be noticed that this short proof relies on the fact that a chord subtends the same angle at all points on either of its arcs as well as on a criterion of similarity based on equality of angles, going beyond the traditional mutually orthogonal pairs of sides. That can be avoided if desired, for example by using the Brahmagupta property of the quadrilateral $ABEC$ (with $ADEC$ as the corresponding pre-Brahmagupta quadrilateral, Figure 8.5). There are other results of Narayana that are most simply derived by working with the same criterion of similarity; as Sarasvati Amma has remarked, whether this is the way he obtained them is unknown ([SA]).

An immediate corollary is a pleasingly symmetric expression for the area of the triangle as the product of the three sides divided by twice the circumdiameter. This result is the main input in many of Narayana's propositions. The particular proposition of interest to us concerns the ratio of the

diagonals and starts by observing that the area $\mathcal{A}$ of the quadrilateral $ABCD$ is the sum of the areas of either of the two pairs of triangles, ABC and ADC or BAD and BCD, into which the diagonals divide it:

$$\mathcal{A} = \frac{AC}{2\mathcal{D}}(AB \cdot BC + AD \cdot CD) = \frac{BD}{2\mathcal{D}}(AB \cdot AD + BC \cdot CD)$$

or,

$$2\mathcal{A}\mathcal{D} = d_1(a_1a_2 + a_3a_4) = d_2(a_1a_4 + a_2a_3).$$

The ratio of the diagonals is thus determined as a function of the sides alone and it is the same as for a Brahmagupta quadrilateral. At this point the third diagonal is not in the picture; this is a result Brahmagupta could have known and used. But we can invoke transposition invariance of the area, $2\mathcal{A}\mathcal{D} = d_3(a_1a_3 + a_2a_4)$, to fix the ratio of the third diagonal to the other two as well:

$$d_1 : d_2 : d_3 = (a_1a_2 + a_3a_4)^{-1} : (a_1a_4 + a_2a_3)^{-1} : (a_1a_3 + a_2a_4)^{-1}.$$

The third diagonal really comes into its own in the most elegant of Narayana's propositions. In his own words (it comes soon after the statement on the existence of the three diagonals),"The product of the three diagonals divided by twice the diameter [is the area]":

$$2\mathcal{A}\mathcal{D} = d_1d_2d_3.$$

As expected, by setting $d_3 = \mathcal{D}$ we recover the corresponding formula for the Brahmagupta quadrilateral.

Before outlining a proof, let us first note how Brahmagupta's theorem on diagonals now follows directly. By equating the two alternative expressions for $2\mathcal{A}\mathcal{D}$, say $d_1d_2d_3 = d_3(a_1a_3 + a_2a_4)$, we get the product formula (Ptolemy's theorem)

$$d_1d_2 = a_1a_3 + a_2a_4.$$

Thus, their ratio being already known, the individual diagonals are determined as functions of the sides, exactly as in the special case of Brahmagupta quadrilaterals. And, as a bonus, we have extensions of the product formula covering the third diagonal

$$d_1d_3 = a_1a_4 + a_2a_3, \qquad d_2d_3 = a_1a_2 + a_3a_4$$

and hence a formula for its length

$$d_3^2 = \frac{(a_1a_2 + a_3a_4)(a_1a_4 + a_2a_3)}{a_1a_3 + a_2a_4},$$

perfectly consistent with the operations of transposition.

The notion of the third diagonal is the capstone on Brahmagupta's approach to the properties of diagonals capturing, in the most economical manner, the consequences of transposition operations. In principle one could do without

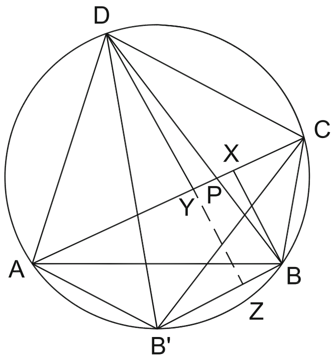

Figure 8.6: Proof of Narayana's theorem: $2\mathcal{A}\mathcal{D} = d_1 d_2 d_3$

it as Brahmagupta likely did in practice but, as Kichenassamy's reconstruction shows, it is laborious work. The formula $2\mathcal{A}\mathcal{D} = d_1 d_2 d_3$ and its proof, which I now give following *Yuktibhāṣā*, are a perfect demonstration of the simplicity it brought to the subject.

In Figure 8.6, $AB'CD$ is the transposed quadrilateral obtained from $ABCD$ by interchanging the sides AB and BC. AC, BD and $B'D$ are the three diagonals of length d_1, d_2 and d_3 and BX and DY are altitudes of the triangles ABC and ADC. Since $B'B$ is parallel to AC, DY extended intersects $B'B$ perpendicularly, at Z. We have then $\mathcal{A} = (BX + DY)AC/2 = (BX + DY)d_1/2$. But $BX + DY = DY + YZ$ is an altitude of the triangle $DB'B$ with the same circumdiameter $\mathcal{D}$ and so is equal to $BD \cdot B'D/\mathcal{D} = d_2 d_3/\mathcal{D}$. Hence $\mathcal{A} = d_1 d_2 d_3/2\mathcal{D}$.

It is easily checked that the proof is independent of the way I have chosen to draw the figure, of whether Z falls between B and C or not (i.e, whether $DB'B$ or DBB' is an obtuse angle), for instance. It really depends on nothing more than the transposition invariance of the area as expressed through the third diagonal. (It can be rewritten with little extra effort by starting from the equality of the areas of triangles ABC and $AB'C$).

Considering that the third diagonal finds its natural habitat in Brahmagupta's whole approach to cyclic quadrilaterals, it is a little bit of a mystery that Narayana, having done all the necessary work, does not take the final easy step and actually state the formulae for the individual diagonals. Nor is there a statement of the area theorem but that can probably be accounted for by the fact that the third diagonal (more generally one of the three diagonals) is associated closely with the diameter and the area formula has no place for it. The fact is that eliminating the diameter from the several formulae for the area that we have met above – equivalently, generalising the steps that work so readily for the Brahmagupta quadrilaterals – is a more demanding task than it is for the diagonals. It is fair to say that the notion of the three diagonals does not have a role in its first documented proof, that given in *Yuktibhāṣā*, but it is also easy to see in it a working out of ideas that go all the way back to

Brahmagupta, 900 years in the past. The following summary of the elaborate *Yuktibhāṣā* proof of the area theorem is very slightly adapted to bring out this point and is best read in parallel with the account of the theorem in the special case of the Brahmagupta quadrilateral given in the previous section.

As we saw there, the key simplifying relation in the special case is the equality of $a_1^2+a_3^2$ and $a_2^2+a_4^2$; so a good starting point is to compute their difference in the general case. Going back to Figure 8.6, let P be the intersection of the first and second diagonals AC and BD. In the special case of the Brahmagupta quadrilateral, P, X and Y coincide and we may think of the length of XY as a measure of the deviation of a cyclic quadrilateral from being Brahmagupta (*Yuktibhāṣā*'s technical term for XY, *laṃbanipātāntaram*, the separation between the feet of the altitudes, conveys the idea). From the right triangles ABX and CBX, we have $a_1^2-a_2^2 = AB^2-BC^2 = AX^2-CX^2 = AC(AX-CX)$ and, from the triangles ADY and CDY, we have $a_3^2-a_4^2 = AC(CY-AY)$. So $a_1^2+a_3^2-a_2^2-a_4^2 = AC(AX-AY+CY-CX) = 2AC\cdot XY$ or,

$$XY = \frac{a_1^2+a_3^2-a_2^2-a_4^2}{2d_1}.$$

This relation also does not depend on the relative positions of X and Y on AC.

The next step is to relate XY to the deviation of the area $\mathcal{A}$ of the quadrilateral from its Brahmagupta value $d_1d_2/2$. The area is the sum of the areas of the triangles ABC and DBC: $\mathcal{A} = d_1(BX+DY)/2$. Now the difference of d_2 and $BX+DY$, like XY, is also a measure of the deviation of the general quadrilateral from being Brahmagupta and the two must be related. Indeed $BX+DY$ and XY are equal to the sides of a rectangle whose diagonal is the second diagonal $BD(=d_2)$ of the quadrilateral (see Figure 8.6):

$$(BX+DY)^2 = d_2^2 - XY^2.$$

This last step is as given in *Yuktibhāṣā*; we can get the same result also from the similarity of the right triangles BXP and DYP, which is perhaps more in the spirit of the area calculation in the special case. The area in terms of XY is thus given by

$$4\mathcal{A}^2 = d_1^2(BX+DY)^2 = d_1^2d_2^2 - d_1^2XY^2 = (d_1d_2+d_1XY)(d_1d_2-d_1XY).$$

Finally, substitute for d_1d_2 from the product formula and for XY from the expression derived above (and supply an extra factor of 4) and we get the formula that we guessed at earlier by symmetrising the area of the Brahmagupta quadrilateral. Its equality with Brahmagupta's $4(s-a_1)(s-a_2)(s-a_3)(s-a_4)$ has already been demonstrated there.

All of this still leaves open the question of how Brahmagupta himself arrived at these theorems. Most of the ingredients of the proofs as described in the late texts on which the account above is based – the *Śulbasūtra* theorem of the diagonal, the properties of similar right triangles, the area formula for the triangle – were standard by the time of Brahmagupta. It is true that neither the

concept of the third diagonal as a means of studying transposition operations, nor of the ‘altitude-separation’ XY as a geometric parameter characterising a general cyclic quadrilateral, is to be found in texts earlier than Narayana’s *Gaṇitakaumudī*. But we have seen in the course of this section that both of these are natural offshoots of notions that Brahmagupta himself introduced, namely transpositions and the Brahmagupta quadrilateral; they, or something which served the same purposes, would have been unavoidable as soon as he went on to the study of general cyclic quadrilaterals. Narayana and Jyeshthadeva were perhaps harking back to knowledge that lived on in an oral transmission chain but which happened not to get put down in a book, an eventuality not unknown in India.

That brings us to the close of Part II of this book. It is by no means a comprehensive account of all the mathematical activity that took place in the classical phase, from the time of Aryabhata to that of Narayana. My attempt, instead, has been to shine a light on the most significant – and the rare paradigm-changing – stages in the development of mathematical thought during the period and, very inadequately, the causal relationships among them and with earlier ideas; similarly inadequate is the account of the historical and cultural backdrop against which they played out. The special position of Aryabhata as the initiator of the resurgence is clear as is that of Brahmagupta for his originality and power allied to a structural and abstract mode of thinking and of Bhaskara II as the consolidator, teacher and propagator of all the new mathematics that came out of it. The stimulus for the renaissance came from two directions, the demands of the discipline of mathematical astronomy, a Greek import unknown earlier in India, and the indigenous traditions of geometry and arithmetic (and the accompanying algebraic mindset), traceable to very early times. And each one of these three great figures exerted a strong influence, each in his own way, on all the mathematics that came after them.

There was a time – which lasted till the early decades of the 20th century – when it was thought by European scholars that mathematics in India breathed its last with Bhaskara II. We have known for some time that that was far from true; Narayana, whose many-sided virtuosity lit up the concluding scene of the classical phase, was still to come. And then there was the last act, a final flowering of the mathematical spirit that found fresh inspiration in Aryabhata’s ideas on infinitesimal trigonometry and turned his vision into the reality of the calculus of trigonometric functions. That story is dominated by the genius of Madhava and is the theme of Part III.

Part III

Madhava and the Invention of Calculus

9

The Nila Phenomenon

9.1 The Nila School Rediscovered

Along the coast of the state of Kerala, roughly half-way down, the river Bharata (Bhārata) as it is called now flows into the Arabian Sea. Its classical name, going back a long way in history – several others were current at various times – is Nila (Niḷā).[1] Though it is only just above 200 km long, in its lower reaches it is very broad, a splendid river even today despite the inevitable degradation. As far back as collective memory and the stories that feed it – half factual, half legendary – go, it has been the backdrop for much of the cultural and intellectual history of Kerala. About the middle of the 13th century, it also acquired a political role when it fell under the rule of the kings of the Zamorin (Sāmūtiri in Malayalam) dynasty of Calicut (Kozhikode; Calicut is the name by which it was always known to the foreign traders), which is not more than 50 km or so to the north. The village of Tirunavaya on the northern bank, about a dozen kilometers upstream from its confluence with the sea, is the site of an ancient temple and it came to be regarded as the source of the secular and sacred legitimacy of the Zamorins. The river and the harbour at its mouth were important for trade as well as militarily. More relevantly as far as the story of mathematics is concerned, villages in the lower basin of the river were among the many settled by Brahmins already before the Zamorins' rise; a few of them had become influential enough to have played a part in establishing

[1]In Malayalam (Malayāḷam), the language of Kerala (Kēraḷa), the consonant *ḷ* (the retroflex *l*, see Chapter 0.4 on the syllabary) is a survivor from the days when it was a local dialect of early Tamil, before its Sanskritisation. (It can also be thought of as the Sanskrit semi-vowel *ḷ* treated as a consonant. Similar is the case with *ṛ*). Another survival from Tamil is the occurrence of both long and short forms of the vowels *e* and *o*: the overbar denotes the long form as in 'Kēraḷa' above (unlike in Sanskrit which has no short *e* or *o*). Throughout this part of the book, in references to the Nila texts, the orthography followed is that found in the particular text, even for Sanskrit words in which there are occasional slight deviations from their classical forms. It is useful to keep in mind that surviving Sanskrit manuscripts in Kerala are almost invariably in Malayalam script.

P. P. Divakaran, *The Mathematics of India*, Sources and Studies in the History of Mathematics and Physical Sciences, https://doi.org/10.1007/978-981-13-1774-3_9

the Zamorins' sway over the region, a measure of how quickly and thoroughly these outsiders from the north had learned to impose themselves on the locals.

It is in some of these Brahmin-dominated villages, under the protection and patronage of the prosperous kings of Calicut, that there emerged a line of astronomer-mathematicians who were to transform, once more and for one last time, the spirit and content of the mathematics of India. By the acclamation of those who followed him in the Nila villages, the founder of the school and the originator of all that was new in their mathematics was Madhava. He is as much of an enigmatic figure as is Aryabhata and in one respect even more so – not one significant mathematical work of his own has come down to us. And, as in the case of Aryabhata, we know nothing definite at all about his immediate mathematical antecedents and very little about his personal life; it is as though an exotic and alien plant took root and flourished among the rice fields and coconut groves of the Nila landscape. We do know however that he left behind a direct line of outstandingly good disciples, several generations of them. Quite a few were prolific authors, compensating to an extent for their master's silence; the mathematics is thus very well documented. Historical references are nevertheless sketchy, scattered allusions to a guru or a patron and the village or the family to which they belonged being our only means of reconstructing the lines of intellectual continuity that linked them and the social and geographical setting in which they lived and worked. Unlike Brahmagupta or Bhaskara II, they may not always say who their father was but they do pay homage to their gurus.

From such circumstantial evidence, historians have been able to piece together an uninterrupted transmission line of teacher to student, from Madhava, born in mid-14th century, to Achyuta (Acyuta Piṣāraṭi) whose death around 1620 marks a natural termination point for the line. (There were a few others who carried the torch until later, down to the beginning of the 19th century, but they added little original to their legacy). As remarkable as the longevity of the line is the fact that all of them were born, and learned and taught their mathematics, in a handful of villages – Trikkandiyur, Alattiyur, Triprangode and Shukapuram – each a short walk from the other (and from the river). Such a concentration of mathematical talent over a long two centuries and a half was unprecedented, in India or elsewhere, and not to be duplicated until the great university towns of Europe took to mathematical scholarship.

The six or seven generations of the mathematical community of the Nila villages represent a coherence of purpose and taste that is also unique. As we shall see in detail later, Madhava himself was directly inspired by Aryabhata's verses on trigonometry, especially the elliptic *Gaṇita* 11 and its arbitrarily small arcs. While expounding and extending Madhava's achievements, those who followed him devoted at the same time a great deal of time to clarifying the meaning and significance of Aryabhata's words, sometimes directly as in the commentaries of Parameshvara and Nilakantha, but also as viewed through the teachings of the two Bhaskaras and a few others. The mathematical personality of Aryabhata is such a constant presence in their writings that Sarasvati Amma

refers to them as the Aryabhata school. More recently, scholars have begun to speak of the Kerala school, but that is to ignore both geography and history: as we have seen (Chapter 8.1), astronomy in the Aryabhata tradition first made its appearance in Kerala five centuries earlier in Mahodayapuram and Kollam, some 100 and 200 km respectively to the south of the Nila. It seems right to call it, as I have done throughout this book, the Nila school after the river that, figuratively speaking, nurtured it (as it did, more generally, the cultural flowering of the region in the 15th - 17th centuries) and in tribute to the singular originality of the mathematics it produced.

The first exposure of European circles to what was special about the work of the Nila school happened not much after the publication of Colebrooke's translations of Bhaskara II and Brahmagupta. In 1832 Charles Whish, who was posted by the East India Company in Kochi, close to the old capital of Mahodayapuram, presented a lecture at the Royal Asiatic Society in London on four manuscripts he had collected in Malabar. Two of them were, we know now, key texts of the Nila corpus, Nilakantha's *Tantrasaṃgraha* (more for the astronomy than for the 'pure' mathematics) and Jyeshthadeva's Malayalam text *Yuktibhāṣā* (the other two, from the 18th and the 19th centuries, are interesting as evidence that the Nila tradition was still alive, even if barely). The lecture was published two years later in the Transactions of the Society.[2] Not only does Whish mention the trigonometric power series – at first sight the most dramatic of the Nila achievements – in his title, but he also goes on to explain that they came out of the same ideas as they did later for Newton, "a system of fluxions" (using Newton's term for the derivative). Whish knew Colebrooke, another Company man and a highly placed one based in Calcutta, and was encouraged by him in his search for mathematical and other manuscripts. Given their find place in a far corner that was almost outside political India, unsuspected till then of a mathematical culture, and given also his own knowledge of Aryabhatan trigonometry from his study of Brahmagupta and Bhaskara and from Samuel Davis' partial translation of *Sūryasiddhānta*, no one was better positioned than Colebrooke to appreciate what these manuscripts represented. Nothing of the sort of recognition that could have been expected seems to have happened. Whish's lecture and its publication made little lasting impact, to the point that the custodian of his library, Hendrik Kern, betrays no awareness of the Nila works while at the same time relying on a commentary by Parameshvara in the preparation of his edition of *Āryabhaṭīya* (1874).

In 1948, a distinguished scholar of traditional sciences, Rama Varma (Maru) Tampuran - the full name identifies him as a royal, a not infrequent convergence of social background and vocation in Kerala – produced, in collaboration with a mathematics teacher with a modern, western ('English') education, Ayyar, Akhilesvara, a fully annotated critical edition of *Yuktibhāṣā* (Part

[2] Charles M. Whish, "On the Hindú Quadrature of the Circle, and the Infinite Series of the Proportion of the Circumference to the Diameter Exhibited in the Four S′ástras, the Tantra Sangraham, Yucti Bháshá, Carana Padhati, and Sadratnamála", Transactions of the Royal Asiatic Society of Great Britain and Ireland, vol. **3** (1834) p. 509.

I, General Mathematics [YB-TA]) in Malayalam, the original language of the text. We have already had many occasions to refer to this work (it is understood that the references to *Yuktibhāṣā* in this book are always to Part I). The following chapters will make it clear why it is *the* indispensable text for a proper understanding of what was path-breaking about the new mathematics of Madhava. Motivations, conceptual inventiveness and technical advances (including proofs) are all given a meticulous and sophisticated treatment in unambiguous Malayalam prose, a far cry from the enigmatic Sanskrit *sūtra*s of earlier masters and even from most contemporaneous writing. Nevertheless, it is still a work in Sanskrit mainstream style: no diagrams and no symbols and equations. Tampuran and Ayyar supply them as well as the occasional illuminating evocation of parallel ideas and themes from other *śāstra*s, the perfect *vyākhyā* in the best Indian manner. Considering the limited readership it was aimed at, theirs was a heroic effort or, perhaps, it was a tribute to the still living tradition of classical scholarship in Kerala. For those who could read Malayalam, and for those others to whom they communicated their excitement, it was a revelation.

The first real consciousness of what quickly came to be called the mathematics of Kerala dates from this re-re-discovery. The long preparation of the Tampuran-Ayyar edition of *Yuktibhāṣā* inspired the first publications in English after Whish on the mathematics of the Nila school. Over a period of time (1946 -1984) C. T. Rajagopal, mathematician at the University of Madras, wrote a number of articles on the subject (in collaboration with several colleagues; there was also some unpublished work by K. Balagangadharan in the late 1940s which he shared with Rajagopal). Unlike Whish's paper, they did not go unnoticed; D. T. Whiteside, the editor of Newton's mathematical papers, refers for example to what he learned from Rajagopal about the sine series and its Nila proof. While this was going on, Sarasvati Amma's work – also prepared at the University of Madras but in the Sanskrit Department, and she could read Malayalam – which later became the book [SA] was being written and it contained, for the first time in English, full proofs following *Yuktibhāṣā* of all the trigonometric series and much other Nila geometry besides. The following decades saw increasing attention being paid, by scholars in India and outside alike, to the body of mathematical knowledge described in the Nila texts. None played a bigger role in promoting this reawakening than K. V. Sarma – equally at home in Malayalam and Sanskrit – who, over a half-century of devoted scholarship, published critical editions of text after text of most of the main personalities in the Nila line. Analytic histories written during this time did not always result in as authentic a portrait as one might wish, however; part of the reason no doubt lay in the inaccesibility of the balanced and finely detailed *Yuktibhāṣā* in the original (and the Tampuran-Ayyar commentary) to the vast majority of interested historians. That handicap has finally been overcome with the publication in 2008 of K. V. Sarma's own English translation, supplemented by explanatory notes in modern notation by Ramasubramanian, Srinivas and Sriram ([YB-S]). Subsequently, Ramasubramanian and Sriram have also produced an English translation of *Tantrasaṃgraha* together with an elaborate commentary

([TaSa-RS]). Those of us who are not at home in Malayalam and Sanskrit have now in hand most of the essential primary material for an objective assessment of the Nila corpus as a whole. One gap remains; there is still no translation of Nilakantha's *Āryabhaṭīyabhāṣya*, particularly valuable as a guide to the line of thought that connects Madhava to Aryabhata.[3]

What then is new in the mathematics of the Nila school? There is an eye-catching first answer: for the first time in history, Madhava wrote down the power series expansions of the trigonometric functions arctangent:

$$\theta = \arctan(t) = t - \frac{t^3}{3} + \frac{t^5}{5} - \cdots,$$

sine and cosine:

$$\sin\theta = \theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \cdots,$$

$$\cos\theta = 1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \cdots$$

as well as, also for the first time, an exact numerical expression for π as an infinite series:

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \cdots.$$

The fact that these series were also among the spectacular early successes of European calculus served to give their discovery by Madhava, almost three centuries earlier, a special cachet among historians (when they finally got round to them). But this benchmarking against a very particular set of series common to both traditions had also the effect that it took the spotlight away from the fact that the breakthrough that led to them was the same as in 17th century Europe, the invention of calculus. And there were other notable advances, modest only in comparison with the precise articulation of the infinitesimal philosophy – the metaphysics of calculus in the words D'Alambert used later in Europe – in which the new mathematics was embedded, that did not get their due. Two examples out of several that stand out are the explicitly stated recognition that infinite series can meaningfully and exactly represent well-defined mathematical quantities, and the identification and the systematic use and justification of mathematical induction to validate an infinite sequence of propositions indexed by the positive integers.[4] These issues will engage us through most of this part of the book, but a preliminary reference to some historiographic points they bring up – especially the continuing discussion and debate on whether Madhava's approach to trigonometric series qualifies as the founding of the discipline that came to be called calculus – may be helpful here as a statement of intent.

[3]A very partial sample, largely non-mathematical, of the riches it contains will be found in N. K. Sundareswaran, *The Contribution of Keḷallūr Nīlakaṇṭha Somayājī to Astronomy*, Publication Division, University of Calicut (Kozhikode 2009).

[4]As a footnote, we may note the chagrin Newton felt on learning that his binomial series for fractional exponents were not the first examples (in Europe) of infinite series, that he had been anticipated by Niklaus Mercator in Germany with his logarithmic series.

It is not a matter of chance that the writing down of the first trigonometric power series in Europe came on the heels of the invention of calculus, at the hands of the two mathematicians credited with its invention, Newton and Leibniz (and Gregory). The arctangent series for general angle is the simplest nontrivial example of the rectification of a curve (an arc of the circle) and the problem could not have been addressed without a clear appreciation of the notions of local linearisation (differentiation), its inverse (integration) and the relationship between the two (the fundamental theorem of calculus). Its derivation in the Nila texts follows the same path; it is in fact a textbook illustration of how indispensable the quantification of infinitesimal change – in the form of differentials, more or less in the spirit of Leibniz, rather than the fluxions of Newton – is for determining the lengths of curves. The problem is posed explicitly as one of rectification of the arc (*cāpīkaraṇam*, 'arcification') and, once expressed in symbols and equations, the *Yuktibhāṣā* proof of the series is the same in all essentials as that taught in a calculus class today. In other words, the conceptual advance that underpinned the Nila discovery of the arctangent and the π series is none other than the fundamental principle of infinitesimal calculus, in the particular form of a reduction to simple quadrature of a rectification problem. The sine and cosine series will be seen, as we work through them later on, to require techniques more demanding than the integration of an explicit function, leading effectively to a method of solving exactly the second order difference (differential in the limit of small angular differences) equation satisfied by the functions. Much of Part III of this book is presented from a standpoint that emphasises the calculus (while remaining faithful to the texts), partly as a necessary corrective to some of the existing writing but, more pertinently, to make the point that Madhava's finest accomplishment was the laying of the foundations of calculus as an independent discipline rather than as the discoverer of some new results in trigonometry, no matter how spectacular.

The conceptual advances are matched by a formidable technical virtuosity in the working out of their consequences. New algebraic and analytical tools, building on the foundations laid by Brahmagupta and Aryabhata but going far beyond them, were deployed. Most remarkable among them, worthy of being taken note of in this overview, is a method of estimating the remainder when the π series is truncated after a finite number of terms and of reordering the series itself for faster convergence. Problems of this nature were not addressed in Europe until well into the 18th century. The basic philosophy relied upon in this 'approximation theory' was an old favourite, that of recursive refining, which we have already encountered in several different forms in this book. The Nila mathematicians thought of the method in very general terms, gave it the generic name *saṃskāram* and made it into an instrument of great flexibilty and power.

Of the enormous and varied output of the Nila school, the work on which the following chapters lean most heavily is Jyeshthadeva's *Yuktibhāṣā*. There are excellent reasons for this, quite apart from the expository qualities alluded to earlier. The task it sets itself at the outset is that of providing an account of

the mathematical material required in the planetary model that Nilakantha had already (25 years earlier) expounded in *Tantrasaṃgraha*. But in actual fact the work transcends this remit. It is broad enough in its content and organisation to serve as a definitive record of what is truly deep in the mathematics of 15th-16th century India, at least as far as the disciplines central to Jyeshthadeva's main interest are concerned. That interest consists primarily, though not exclusively, of the founding principles of the infinitesimal calculus of Madhava in their application to trigonometric functions, along with all the supporting structures they called for. To these virtues of thematic clarity and a fine mathematical taste, add its superb logical organisation and the carefully composed and presented proofs of virtually all significant results cited, old or new, and we have a book that bears comparison with the best monographs of modern times, never mind the absence of figures and equations.[5] The other work whose insights often throw an illuminating sidelight on how the Nila school understood Aryabhata's ideas is of course Nilakantha's *Āryabhaṭīyabhāṣya*. As we shall see below, teacher and disciple probably wrote their two masterpieces, each in his own very distinctive style, within a couple of years of each other, when Nilakantha was a wise old man in his late seventies or early eighties. It is an intriguing thought, for what it is worth, that both of these are prose works.

9.2 Mathematicians in their Villages – and in their Words

Given the general paucity of information regarding lines of transmission of mathematical knowledge – not to mention personal and biographical details – in India, it is something of a welcome surprise that we can build up a fair though still incomplete picture of mathematical life in the Nila villages. For this, we have to be grateful to the numerous books they produced and to Kerala's unbroken tradition of astronomical theory, until recent times, (and astrological practice, popular to this day) that ensured their survival. Texts of the Nila authors as well as those from other parts of India were copied and recopied on palm leaf (and on paper in later times). New *bhāṣya*s and *vyākhyā*s continued to be written; a very elaborate commentary in Malayalam on *Līlāvatī* by a traditional astronomer-astrologer was published as recently as 2007.

The most generous among the Nila authors in matters of historical and biographical interest was Nilakantha, primarily in the more discursive passages of his commentary on the *Āryabhaṭīya*, and it is no surprise that he is the only one whose birth date (in 1444) is precisely known, from his own words; consequently, it has become the pivotal date around which every attempt at the reconstruction of the Nila chronology has to revolve. Nilakantha also tells

[5]Some aspects of the difficulties involved in carrying out such a technically demanding project – without the aid of symbolic notation – and their resolution are brought out in P. P. Divakaran, "The First Textbook of Calculus: *Yuktibhāṣā*", J. Indian Philosophy, vol. **35** (2007), p.417.

us that his teachers were Damodara (in astronomy) and Ravi (in more philosophical and spiritual matters), sons of Parameshvara, and that Parameshvara himself, in whose house he spent time as a young resident student, was the disciple of Madhava. That is four overlapping working lives covered, from Madhava's birth (roughly 1350) to Nilakantha's death (around 1530). Those who came after Nilakantha say little or nothing about themselves – and no grandson put up a memorial inscription, tracing the family glory five generations back, for any of them. But there are incidental references to gurus and other venerated elders and to their villages or family affiliations which have been put to good use in the pioneering work of K. K. Raja and, independently, of K. V. Sarma[6] to arrive at a reliable sequence of teachers and their pupils as well as, in a few cases of doubt, to assign texts to the correct authors.

But putting dates to their lives and, especially, to their principal works, except approximately, is another matter. (Among the very few exceptions is Nilakantha's *Tantrasaṃgraha* which is internally dated). Occasional help is available in the form of a recorded date, for example as the initial epoch ($t = 0$) in an astronomical computation occurring in a particular manual, or as a reference to a respected man of learning by a contemporary whose date is known. But there are risks in putting our faith in the casual words of supposed contemporaries; nothing illustrates them better than the fact that the only way to reconcile every such reference is, for instance, for Parameshvara, Damodara, Nilakantha and Jyeshthadeva all to have been centenarians and been intellectually productive till the end of their lives (see K. V. Sarma, cited above). That is of course not impossible but is not required by anything else we know about their lives and work.

One reason for being cautious about giving equal credence to every unsupported fragment of information has to do with the common custom in Kerala (as in some other societies) of designating the senior male in a family by the name of the family alone, the 'house': a reference to a 'Kelallur' in the middle of the 16th century may just mean the patriarch of the Kelallur house at that time, not necessarily *our* Kelallur Nilakantha. There are other areas of potential confusion. When the 'place' to which a person belongs is added to his name, the place meant may be his house (*illam* or *mana* in the case of Brahmins) or his village. House and village names are often given in two variations, the original local (Malayalam) name and a Sanskritised variant, not always formed in a systematic way. And, to confound matters further, names like Narayana, Parameshvara, Nilakantha and Shankara were (and are) extremely common among Kerala Brahmins and so of doubtful value by themselves as unique identifiers. It may seem that these cannot be serious difficulties but, in the context of a general lack of trustworthy historical records, one can sometimes be misled by them, as we shall see.

---

[6]K. Kunjunni Raja, *Astronomy and Mathematics in Kerala*, Adyar Library and Research Centre, Madras (1995) (first published in the Adyar Library Bulletin in 1963); K. V. Sarma, [Sa-HKSHA] and the Introductions (in English) to several of the Nila texts edited by him.

The following brief account of the lives of the chief protagonists of the Nila school is based largely on the investigations of Raja and Sarma cited above, supplemented by what can be dredged from remembered local histories and in the light of a more critical reevaluation of the evidence. It is necessarily provisional: we can always hope for some fresh discovery that will lead to a less fuzzy picture than we have at present. Also, Madhava himself is excluded from this collective biographical sketch. His case is somewhat special: so little is known about his life and mathematical background, so much of what has been written about him is without a real foundation and, at the same time, so critically important is he to the whole Nila story, that a fresh look, even if partially speculative, would seem to be justified. As in the very similar case of Aryabhata – who was Aryabhata? – that is best done in a later section, as part of the pre-Nila evolution of mathematics in Kerala and elsewhere and its historical backdrop.

The only disciple of Madhava that we know of was Parameshvara, a very fine astronomer, both mathematical and observational. Among his contributions was a revision of the planetary parameters central to the Aryabhatan system (over an already existing revision due to Haridatta dating from 683), described in a work finished in 1431 and named *Dṛggaṇita*. The title can be loosely translated as "computations based on observations" and the book was the result, according to himself (and Nilakantha), of fiftyfive years of observations of heavenly bodies in the skies above the sands of the northern bank of the Nila. If we assume that he embarked on his life as an astronomer while still a *brahmacārin*, a young scholar-in-training, he would have been born around 1360 (this is in fact the main support for placing Madhava's birth in the middle of the century). Another of Parameshvara's books is dated 1443, a year before Nilakantha's birth, but a better estimate of his longevity is provided by the latter's recollection of the time he spent in the old guru's home. We can conclude with a degree of confidence that Parameshvara did have a long life, well into his nineties.

The name of his *illam* was Vatasseri (Vaṭaśśēri) or Vaṭaśreṇi – an example of Sanskritisation devoid of semantic justification, possibly based on phonetic similarity alone, the rebus principle in action – which he places on the northern bank of the Nila, at its confluence (*saṃgama*) with the ocean. He also says that he belonged to Aśvatthagrāma, identified with the village of Ālattiyūr, and that is a good semantic fit, both meaning 'the village of ficus trees'. This relative wealth of information nevertheless masks several ambiguities. Modern Alattiyur is certainly the Alattiyur of the 15th century, identified, at both times, by a temple complex that is much older – and temples are not movable objects. But Alattiyur is not where the Nila flows into the ocean nor even particularly close to the river or the sea. As for the house name, it has been suggested that the Sanskrit *vaṭaśreṇi* is synonymous with the Malayalam name Alattiyur and it is a fact that *cēri* in Vatasseri also means a village, not a house. Add to this the perfectly acceptable sense of *vaṭa* in Malayalam as 'north' (Vatasseri = the village to the north), and it would seem that everyone is free to pick

and choose among the various combinations of meanings and identities. The consensus nevertheless is that Parameshvara passed his long and active life in a house called Vatasseri in or near the village of Alattiyur.[7] There is no compelling reason to question it; as Jyeshthadeva says in *Yuktibhāṣā* (in a very different context), one must have something definite to anchor oneself to.

Parameshvara was a loyal Aryabhatan and is the author of well over twenty books, covering a wide variey of topics and types: original monographs in astronomy and related mathematics, manuals of computation (of eclipses especially), erudite commentaries on most of the older classics with an emphasis on Bhaskara I, and even books on astrology. Trigonometry and the sine table figure prominently though not Madhava's infinitesimal extensions of them. Unexpectedly, in his books on the sphere and the circle there are some results in the new circle geometry in the style of Narayana Pandita's propositions in *Gaṇitakaumudī*, which was completed in 1356, about the time of his birth or just before. What makes this conjunction of dates and mathematical themes intriguing at first sight is that Nilakantha mentions a Narayana, son of (another) Parameshvara, as one of his teacher's teachers (along, of course, with Madhava). As it happens, this particular instance of nominal degeneracy is resolved by their fathers' names; the father of Narayana Pandita was Nṛsimha – another wishful idea runs aground and, in a way, it is a pity. As Sarasvati Amma was the first to remark, there are many commonalities between the work of Narayana of *Gaṇitakaumudī* and some of the Nila mathematics. I will try to add some speculative substance to his shadowy life later on (section 4 below).

Little is known about Damodara apart from the high esteem in which he was held by Nilakantha. A credible estimate of the period of his life will nevertheless turn out to be indispensable in fixing the dates of some of those who came after, Jyeshthadeva in particular.

Nilakantha (Nīlakaṇṭha Somayājin, the honorific indicating that he was an adept of Vedic ritual performances) was born into a house called, in his time, Kelallur *mana* in Trikkandiyur (Tṛkkaṇḍiyūr) village, about 6 km north of Alattiyur. As in the case of Parameshvara, the name of the house was transformed, first into Kerala-nallur, for no good reason one can find, and then into its Sanskrit equivalent Kerala-sadgrama, the 'good village of Kerala'. These mystifying name changes – it is very unlikely that the citizens of the Zamorin's small kingdom thought of themselves as inhabitants of a greater Kerala – and the vagaries of the popular narrative they represent are a persistent irritant in any attempt at reconstructing a reasonably accurate history. As far as Nilakantha is concerned, it may not be a serious issue; his house or at least the compound in which it stood still exists, along with the open water tank said by the local community to have been used by him, close to the main temple of the village – it helps that Trikkandiyur has a long and continuous history, partially set down in the records of the temple. There is nevertheless a good

[7]Some confusion can be avoided by noting that Alattiyur is not the same as the bigger town of Alattur, some hundred kilometers to the east, though the locals often pronounce it that way.

reason to dwell on what may seem to be an inconsequential matter: we will encounter these irritants again when we turn our attention to Jyeshthadeva and Madhava and, in Madhava's case, a satisfactory resolution of similar locational ambiguities may well turn out to be of help where we badly need some: who was he and what was his mathematical background?

Nilakantha led a full life. Apart from mathematical and observational astronomy and the obligatory Sanskrit grammar (and, presumably, the conduct of rituals), he was well versed in various philosophies. He travelled widely within Kerala, spreading the word about the superiority of Parameshvara's newly established planetary parameters, and was known to scholars outside its boundaries; among the Nila stalwarts and along with the later Achyuta, his is the name best remembered in the cultural history of central Kerala, woven into the very fabric of folk memory as it were.

In the midst of this busy life, Nilakantha found time to write several books, at least eleven of them according to K. V. Sarma, ranging from the brief and technical to the discursive fount of wisdom that is the *Āryabhaṭīyabhāṣya*. Of them, three or four are worthy of special notice. The best known is, of course, *Tantrasaṃgraha*, which can fairly claim to be the *ur*text of the Nila astronomical doctrine in its final form, in 432 *sūtra* style two-line verses. (An English translation accompanied by detailed annotations and the original Sanskrit text has recently been published [TaSa-RS]). Infinitesimal trigonometry is largely absent. One can see why the need to augment it with the new mathematics of Madhava and a much more detailed working out of the astronomy was felt by Jyeshthadeva and Shankara.[8]. The work is dated by two verses of benediction, its first and last, which, according to Shankara, have a second sense as *kaṭapayādi* counts of the *kali* days on which it was begun and finished. The astonishing fact is that they are separated by just five days – five days to complete a highly technical monograph in metrically rigid Sanskrit verse. Anyhow, the dates fall in March 1500, almost exactly a thousand years after the *Āryabhaṭīya* was composed and two years to the month after the landing of the Portuguese sea-captain Vasco da Gama on a beach north of Calicut, an event that was to have a momentous impact on mathematics in the Nila basin and, eventually, in India as a whole.

Three books that Nilakantha wrote late in his life – as attested by him or by the internal evidence in them – share a special affinity. All three go well beyond the mathematical and astronomical themes which are their ostensible concern to hold a mirror to the epistemic foundations of the practice of astronomy and mathematcs and the historical context of Nilakantha's own personal and professional life, all with a pronounced post-Madhava slant. Of these three wise books, one is a *vyākhyā* of his own earlier *Siddhāntadarpaṇa* ("A Mirror to *Siddhānta*") and it is here that we find, finally, a grudging acceptance of Aryabhata's rotating earth: that, since motion is relative, it does

[8]There used to be some confusion, going back to the manuscripts relied upon in [YB-TA], about what was originally in *Tantrasaṃgraha* and what was interpolated into some of them from the verse commentary of Shankara. That seems to have been sorted out now.

not matter in astronomical computation whether the celestial sphere is moving from east to west or the earth is spinning from west to east, thus turning the old master's words against himself. Among other gems, it also has a delightful discourse on mathematical truth and its verification, with the theorem of the diagonal as a case-study. Such epistemic preoccupations take centre stage in *Jyotirmīmāṃsā* ("Investigations in Astronomy"), effectively an essay on the nature of scientific endeavour, on the means by which knowledge is to be acquired and validated as exemplified specifically by astronomical phenomena and their description. It is probably the first time in India that such an analysis was made by an active scientist, outside the abstract theorising of linguists and logicians and metaphysicians. To respect for the teachers of the past – and while firmly rejecting divine revelation – Nilakantha adds the theoretical criterion of sound inference or informed conjecture and, perhaps to the surprise of some, the imperative need to update the knowledge base by continuing observations as the basis of all knowledge; the whole process is to be consolidated through open discussion and instruction so as to continually create and safeguard a living tradition of science. This would have been a radical position for any astronomer (Aryabhata himself excepted) to take but coming from Nilakantha, who was a model of orthodoxy in his observances, it was almost heretic, reminiscent of the materialistic *lokāyata* system of philosophy, a school of thought generally held to be beyond the pale. Altogether, the book is a passionate manifesto for the vocation of an astronomer-mathematician-pedagogue by one who has practised it himself and thought deeply about its philosophical moorings.

The third of the wise books, the *bhāṣya* of the *Āryabhaṭīya*, was in fact finished before the other two – there are references to it in them – though not much before; Nilakantha himself says that he was of advanced age when it was completed. The style and the insistence on epistemic soundness, as well as more concrete pieces of textual evidence, all suggest that the three works are roughly contemporaneous. Indeed it is quite evident that the general principles of scientific enquiry that he lays down in *Jyotirmīmāṃsā* and the *vyākhyā* on *Siddhāntadarpaṇa* are the distillation of the method he adopted in the decoding of Aryabhata's aphoristic verses – or, conversely, that what we see in the *Āryabhaṭīyabhāṣya* is the practical deployment of an epistemic framework he erected for himself over a lifetime spent in the pursuit of knowledge. It is a long, rich and profound work – and it certainly could not have been written in five days – with insights not to be found anywhere else in the history of mathematics in India. Intellectual resources from the vast store at his command – including the writings of Pingala and Bhartrhari from among those we have met in this book – are brought to bear on one subtle point or the other. Nilakantha was conscious that the work was something special; the text refers to itself as the Great Commentary, *Mahābhāṣya*: if the *Āryabhaṭīya* was the *Aṣṭādhyāyī* of mathematics, Nilakantha was its worthy Patanjali.

As far as the mathematics is concerned, we have already come across instances of Nilakantha's concern for spelling out the logic behind the statements

of the *Gaṇitapāda* of the *Āryabhaṭīya*: his deconstructive ('architectural') demonstrations of the formulae for the sums of sums of natural numbers and the sums of their squares and cubes (Chapter 7.1) and, especially, the derivation of the exact sine and cosine differences and the statement and proof of the exact form of Aryabhata's sine difference rule (Chapter 7.4). The careful and rigorously worked out expositions are typical of his approach to all of Aryabhata's mathematics. Even standard elementary results like the theorem of the diagonal and the expression for the area of the circle did not escape a critical reexamination: the former is given yet another new proof of which the one in *Yuktibhāṣā* is a variant (see Chapter 2.2; it is repeated in *Siddhāntadarpaṇavyākhyā* and still another variant is given in another of his books) and the infinitesimal proof of the latter given in *Yuktibhāṣā* (see Chapter 7.5) is, again, similar to Nilakantha's. One aspect of his mathematical personality that comes across clearly in all this is his predilection for the geometrisation of arithmetical and algebraic propositions and a preference for visualisable demonstrations. Whether these demonstrations are faithful to Aryabhata we have, of course, no sure way of knowing.

At a more general level, it is reasonable and natural to subsume Nilakantha's quest for good proofs in his concern with the method of science. The question of proofs in Indian mathematics is a vexed one, at least among some historians. There used to be a belief in these circles, often expressed strongly, that proofs were not considered necessary – and their logical function not grasped – in India. To anyone familiar with texts in mathematics (not to mention logic), there can be little doubt that belief in the validity of a proposition was never, from the time of the *Śulbasūtra*, a matter of revelation, faith and authority alone (supported perhaps by some examples) but had to be buttressed by a process of reasoning, generally called *upapatti* from at least the time of Bhaskara II. Nilakantha's insistence on clarity and logical order gave this process rigour and a structure, very well illustrated by his own approach to Aryabhata and even better in the formally organised proofs, *yukti* (a word he uses freely), which find their full expression in *Yuktibhāṣā*. In this respect as well, the *Āryabhaṭīyabhāṣya* initiated a decisive turning point, a metamathematical one this time. I will have more to say about *upapatti* and *yukti* in due course.

The most spectacular of Nilakantha's mathematical insights is however one for which he offers no proof: the statement that the circumference and the diameter of a circle are incommensurable, i.e., that π is an irrational number. It occurs in the commentary on *Gaṇita* 10, the verse in which Aryabhata gives his proximate (*āsanna*) value of π. Since the *Āryabhaṭīyabhāṣya* is a work that delights in explanations and demonstrations, we have to suppose that Nilakantha did not have a justification for the claim satisfactory to himself. Here are his words (the italics are mine):

> Why is the *āsanna* value given and not the true value? I explain. Because the true value cannot be given. Why? A unit by which the diameter can be measured without a remainder will leave a remainder

> when used to measure the circumference. A unit which measures the circumference without a remainder will leave a remainder when used to measure the diameter. The same unit used to measure both will never be without a remainder [in at least one of them]. *Greater effort will result only in smallness of remainders, never in the absence of remainders.* That is how it is.

The reiteration of the key idea is a measure of the importance Nilakantha attached to getting it across, a pedagogic device *Yuktibhāṣā* also uses to good effect. There is no room for any ambiguity (the italicised sentence) and the whole passage is in keeping with his high standards of lucid logical exposition. It is also an excellent illustration of the principle of reasoning called *anumāna*, informed inference, not established truth: a conjecture as we would call it today. But as striking as the irrationality conjecture itself is the fact that the issue was addressed at all. Indian mathematics spent two millennia and a half, after $\sqrt{2}$ and π made their first documented appearance in the *Śulbasūtra*, resolutely turning its back on the distinction – it is not even clear that the distinction was made – between rationals and irrationals. What made Nilakantha raise the issue, out of the blue as it were? (Strikingly, he does not restate his speculation that $\sqrt{n}$ cannot be determined exactly (see Chapter 7.1) as the incommensurability of the diagonal and the side of the square, for example). One explanation that offers itself is that Madhava's series for π, and infinite series in general, had reopened these long-ignored matters and had convinced Nilakantha that π could not have a finitely expressed numerical value. Indeed, his commentary on *Gaṇita* 10 also invokes Madhava's 11th decimal-place value of π obtained by truncating the series and estimating the resulting error, introduced with the phrase "Madhava born in Sangamagrama has given an exremely close (*atyāsanna*) value of the circumference" – *atyāsanna* it may be, but it is still *āsanna*.

But it is not as though Nilakantha did not know that an infinite series of rationals could sum up to a rational number. The *Āryabhaṭīyabhāṣya* has in fact a beautiful discussion of how the method of iterated refining, when carried out *ad infinitum*, will lead to the formula for the sum of a convergent infinite geometric series, and it occurs in connection with the π series in whose derivation it is an essential intermediate step. It may well be that the irrationality conjecture is the insight that emerged out of the frustration of trying to express the truncation error of the π series as a rational function of the point of truncation (see Chapter 13 below); Nilakantha's juxtaposition of Madhava's "very close" value of π with its conjectured irrationality will then get some motivational backing. Regrettably, none of those who followed Nilakantha took up the matter. (One of the reasons for which I date *Yuktibhāṣā* before *Āryabhaṭīyabhāṣya* is the lack of any mention of the irrationality issue in the many-sided investigation of π in the former). Of all the grey areas of the Indian mathematical-historical landscape, the background to the irrationality conjecture remains among the more tantalising.

Madhava and the spirit of his mathematics are ubiquitous in the *Āryabhaṭīyabhāṣya*, particularly so in the annotations of the *Gaṇitapāda*; in the long colophon at its conclusion there occurs, after a listing of the virtues of his own '*Mahābhāṣya*', the phrase *mādhavādigaṇitajñācārya-kṛta-yukti* ("proofs constructed by mathematician-teachers such as Madhava"). Many cited verses, encapsulating calculus-related results, are attributed to him. We are left to draw the inference that Nilakantha saw the new infinitesimal trigonometry – calculus on the circle in other words – as the direct outgrowth of Aryabhata's ideas, a belated vindication of his enigmatic vision.

Surprisingly for a work so rich in all kinds of incidental comments, by an author who was the most culturally aware of the mathematical greats, the *Āryabhaṭīyabhāṣya* does not carry a date, nor even any circumstantial indication beyond the mention of the author's advanced years. Neither do the other two wise books, except that they both refer to the *bhāṣya*. Nilakantha was past eighty by 1525. The only reason why historians have been persuaded to prolong his life beyond 1530, say, would be to accommodate a mention of a 'Kelanallur' by an obscure astrologer in 1542. It is not a very strong reason. To assume that the name is a misspelling of Kelallur is an extrapolation which may well be off the mark; even if it is not, it need only mean (as noted earlier) that the Kelallur family was not defunct at that time. Discounting this isolated reference has no impact on the general Nila chronology and it is perfectly reasonable then to assign a date in the mid-1520s to the completion of the *bhāṣya*, leaving the old lion a few more years to ruminate and produce his last two philosophical essays. A slightly compressed time frame fits in better with what we know of Jyeshthadeva's life and the date of *Yuktibhāṣā*, to which I turn now.

At the time of publication of the Tampuran-Ayyar edition of *Yuktibhāṣā*, its authorship was unknown. The subsequent definite identification of the author as Jyeshthadeva is the result of some fine detective work by Raja and Sarma (cited above). The fragments of information unearthed by them also suggest a plausible time frame for his life that is not, I will argue, the same as what they themselves have put forward. These snippets come from two or three passing allusions in the writings of others – there is nothing remotely personal or anecdotal in *Yuktibhāṣā* itself – which we have no reason to question; two of them are in fact as authentic as they can be. Besides, they hold together well without stretching credulity.

The most informative (though possibly not as impeccably authentic as the other two) of these allusions is a passage in mixed Malayalam and Sanskrit at the end of a manusript otherwise not important (see K. V. Sarma's Introduction in English to his edition of *Tantrasaṃgraha* [TaSa-Sa]). It traces the Nila lineage from Parameshvara to Achyuta saying, in passing: ". . . Damodaran's disciple Jyeshthadevan. He was [a] *nampūri* (Brahmin) of (from) Paṛaṅṅōḍ. He was also the person who wrote the book *Yuktibhāṣā*. Jyeshthadevan's disciple Trikkandiyur Achyuta Pisharati". The other two references are from books by Jyeshthadeva's own disciples, in the form of tributes to their guru: Shankara's

Yuktidīpikā has a colophon at the end of every chapter, identical, in which occurs the phrase "the venerable twice-born (Brahmin), resident of Parakroḍa"; and Achyuta (in a text which we will have no reason to refer to again) speaks of his "*sadguru* . . . Jyeshthadeva of advanced age". We may conclude in particular that Jyeshthadeva was alive when his pupils recorded their homage to him.

The information that Jyeshthadeva began his studies under Damodara and then continued with Nilakantha would suggest that he was perhaps a generation or so younger than the latter, born around 1465-75. Formal initiation into the vocation of learning takes place traditionally around the age of 10. By 1470, Nilakantha would have been 25 years old and Damodara about 75 – allowing a 30 year generation gap between him and his father Parameshvara – and that matches their roles, just about in Damodara's case, in Jyeshthadeva's apprenticeship. How long did he live? Sarma notes that, according to a local belief gathered by Whish in the early 19th century (recorded in his Asiatic Society lecture), he was also the author of a dated (1608) text in Malayalam verse, partly narrative, called *Dṛkkaraṇa*. That is possible, even stretching things, only if he was born not earlier than about 1500 (Sarma suggests 1500 - 1610 as his life span). Such a late date cannot be squared with his having been a pupil of Damodara unless Damodara too lived well beyond the age of 100, 1400 - 1520 say(!) and remained vigorous till the end of his life. The problem clearly is with the attribution of *Dṛkkaraṇa* and the obvious solution is to take it out of the equation, removing in one go the constraints on the year of his birth and his death. There are also stylistic arguments to support such a position. The Malayalam prose of *Yuktibhāṣā* is colloquial and very rough-and-ready (sometimes characterised misleadingly as 'medieval') with little attention paid to linguistic and literary niceties – wrong tenses and case endings and incomplete sentences are quite common – while the later text is in the limpid and correct verse style of the time, difficult to reconcile with a common authorship.

But we are still not done with doubts about how long Jyeshthadeva lived. One phrase in Achyuta's homage to his aged guru was read in a later commentary as a *kaṭapayādi* number (1,714,262) and declared to have been the *kali* date on which it was written, which falls in the year 1592 (cited by both Raja and Sarma in support of the later dates). If the reading is authentic, Jyeshthadeva was alive in 1592 and Damodara would still need to have been a near centenarian to have taught him. So the value of this bit of circumstantial support for a late chronolgy depends on whether the particular phrase of seven syllables was intended by the author to be a coded number also. Achyuta gives no indication that he meant it to be a number. Now, any sequential set of syllables can be declared by anyone to have been deliberately meant to be a hidden number; given the flexibility of the coding, and the fact that the first four syllables (since *kaṭapayādi* is read backwards) can be ignored if we are satisfied with a purported date within a latitude of 25 or 30 years, it is not so very hard to find in a not-too-short piece of writing a sequence of syllables

for any three digit number (171 in the present case), corresponding to a date within any interval of 25 years in the author's lifetime.[9] It seems best, once again, to disregard such potentially spurious interpretations of isolated data in favour of a coherent picture that accommodates all the well established facts and reasonable extrapolations therefrom; dating Jyeshthadeva's birth to 1470 or thereabouts contradicts nothing that is securely known. We still do not know when he died but that is not so important, the only constraint being that his pupil Achyuta died in 1621; that can easily be fitted into the general chronolgy without making everyone a centenarian.

There is room for some uncertainty about Jyeshthadeva's native village but that too is only of minor importance. The reference to "Paṛaṅṅōṭṭu naṃpūri" is generally taken to mean that Paṛaṅṅōḍ (the change from *ōḍ* to *ōṭṭu* is just the appropriate case ending) was his *illam*, supposedly in Alattiyur. Grammatically and in common usage, it can just as well be his village, as I mentioned in the beginning of this section. There is in fact a Paṛaṅgoḍ village in the neighbourhood, namely Trprangod with its own ancient temple. Prefixing *tṛ* or *tiru*, meaning 'hallowed', to the name of a temple village is a frequent practice in Kerala – Trkkandiyur and Tirunavaya (and numerous others) in the Nila basin – and the change *ṅṅ* → *ṅg* is also common in Malayalam names. Whether Jyeshthadeva worshipped the god Rama in Alattiyur or Siva three kilometers away presumably had no effect on his mathematics.

Discounting the problematic *Dṛkkaraṇa*, Jyeshthadeva wrote only one book that we know of, *Yuktibhāṣā*. Strictly objectively, we are almost totally in the dark about its date. Speaking only of Part I, the universal habit of attributing amost every original advance in the mathematics to Madhava makes it difficult for the historian to place it in relation to what went before and what came after. Thus, though we know that Shankara's *Yuktidīpikā* is a later work, we cannot tell it from the texts, with their emphases on somewhat different facets of Madhava's calculus, since they are both claimed to be accounts of Madhava's own work. Perhaps a careful analytic comparison of *Yuktibhāṣā* with Nilakantha's *Āryabhaṭīyabhāṣya* – given their epistemic and foundational affinities – will reveal nuances and small details which naturally belong in them but are absent in one or the other. One example concerns the irrationality question: in *Yuktibhāṣā*'s exhaustive treatment of the π series, there is no mention of Nilakantha's conjecture. Is that a sufficient ground to conclude that *Yuktibhāṣā* was written before the *Āryabhaṭīyabhāṣya*? For the time being, it seems prudent just to think of them as roughly contemporaneous, sometime in the 1520s, in the hope that more conclusive evidence will surface one day.

One cannot do justice to the riches of *Yuktibhāṣā* in a short survey such as the present section is intended to be. Most of the rest of this book is in any case about what we can learn from it. I will provide a brief introduction to its

[9] It is an instructive but tedious exercise to search for random occurrences of an arbitrary fixed sequence of numerals in one's favourite text. Is it possible that the implausibly short five days in which the *Tantrasaṃgraha* was supposedly completed is also a result of reading numerical meaning into a piece of plaintext?

content and style in the beginning of the next chapter, more as an appetiser than as a summary; a full and proper appreciation of its singular significance in our understanding of the Nila phenomenon is not possible without going through the mathematics in detail.

Of the major figures still left of the main Nila line, Chitrabhanu (Citrabhānu) was a contemporary of Jyeshthadeva, very likely a student of Nilakantha and resident of a village some distance to the south of the Nila. He is best known for his work on the solution of a class of simultaneous equations in two unknowns, of degree up to cubic, both algebraically and geometrically. The results are not earth-shaking but he was highly regarded, as we learn from Shankara who speaks of his expertise in the *yukti* of arithmetic and algebra (*gaṇita*) and of geometry (*gola*). As we can expect from the stress on *yukti*, the demonstrations are logically clear and and well structured.

Next in order of seniority is Shankara. He was from a village on the river, but well upstream of where all the action has been so far and where he started on his career as a student of Jyeshthadeva. His dates are not precisely known but can be guessed at in relation to his teacher's. The consensus that his productive period spanned the middle decades of the 16th century fits in very well with our early chronolgy for Jyeshthadeva: *Yuktidīpikā*, the earliest of his books, would be justified in addressing Jyeshthadeva as the "venerable twice-born, resident of Parakroda" if it was written around 1540-50 but not if the latter was born in 1500. This and his *Laghuvivṛti* are both commentaries, so described, on *Tantrasaṃgraha*. Late in life he wrote another commentary called *Kriyākramakarī*, ostensibly on *Līlāvatī*, which remained unfinished. (It was subsequently completed by yet another Narayana). The first and the last are extensive works but, as far as the light they throw on Nila mathematics is concerned, it is sufficient to pay attention to *Yuktidīpikā*; in fact *Kriyākramakarī* has large sections bodily lifted from it.

Yuktidīpikā is a *tour de force* of mathematical versification, in the accepted *vyākhyā* style, taking up a verse (or a set of verses) from *Tantrasaṃgraha* for elaboration at great length, ranging over everything he thought might have something to do with it: the 80 stanzas of the second chapter get 1102 stanzas of analysis, for instance. Madhava is placed front and centre; indeed many of the commentatorial verses dealing with trigonometric series were at one time thought erroneously to have formed part of *Tantrasaṃgraha* proper.

Though *Yuktidīpikā* acknowledges handsomely its indebtedness to *Yuktibhāṣā* in the chapter-closing colophons – "Thus have I set out . . . the exposition that has been well-stated by the revered Brahmin of Parakroda" (K. V. Sarma's translation) – the two books are markedly different in their perspective and in the relative prominence given to the various elements, conceptual, technical and computational, that make up the totality of Madhava's body of work. The illuminating asides on the metaphysics of the infinitesimal method and the meticulous explanations of its implementation such as we find in profusion in *Yuktibhāṣā* are largely absent in Shankara. What we find instead are tediously detailed accounts of topics to which Jyeshthadeva barely accords a

nod. A prime example is an interpolation formula which Madhava used to determine sines at values of angles lying between canonical multiples of $\pi/48$. We shall see later that its utility is purely computational; pushing the interpolation to higher and higher orders as Shankara does (but Jyeshthadeva does not) is unnecessary for practical astronomy and, theoretically, leads only to a dead end. It may be that Shankara's writing style was inhibited by the constraints of versification but that cannot be the whole story – 1102 stanzas is a lot of verse after all. The lingering impression is that, for all his diligence and technical strengths, his work suffered from a certain lack of mathematical discernment.

From a historiographic perspective, however, Shankara's work can be said to have had a disproportionately big impact. It came to be known earlier (even before the question of authorship was settled) and much better to historians who read Sanskrit than the Malayalam text of *Yuktibhāṣā* and, to that extent, monopolised their attention: the focus became the end results, to the detriment of what led to them. The result has been to devalue somewhat the true worth of Madhava's breakthroughs.

If anything, *Kriyākramakarī* only reinforces the feeling of Shankara's inadequacies as an interpreter of Madhava. When Nilakantha looked for links with the past, it was to Aryabhata that he turned. Shankara's choice of *Līlāvatī* as a vehicle to convey the new Nila mathematics – or, perhaps, *vice versa* – is itself strange: there is hardly anything infinitesimal in *Līlāvatī*. If Shankara was seeking the source of Madhava's infinitesimal inspiration in Bhaskara II, the fact is that it cannot be found there.

Achyuta, Jyeshthadeva's other and more junior disciple, was a competent astronomer and mathematician and the author of about a dozen monographs and commentaries. He belonged to Nilakantha's natal village Trikkandiyur and, true to the heritage of his illustrious predecessor, was a highly respected public intellectual of his time, an esteemed grammarian of Sanskrit, a poet and a physician and scholar of the traditional system of medicine (*āyurveda*; the Nila basin has always been and still is a stronghold of the 'science of life'). Among his disciples was the famous devotional poet and Sanskrit scholar Melpattur Narayana Bhattatiri whose memorial verse for him lets us place his death in 1621. It is not known when he was born but the date of his death is consistent with our early dating of his guru Jyeshthadeva.

To summarise finally what we have learned from all this chronological hair-splitting: Madhava was born in mid-14th century in a village whose name was Sanskritised as Sangamagrama – Nilakantha says so, more than once: *saṃgramagrāma-ja mādhava*; time of death unknown, but probably not much after 1420 and definitely before Nilakantha's initiation into learning. His disciple Parameshvara, the only one we know of, lived from 1365 ± 5 until at least 1455 and Parameshvara's son Damodara, the key to the whole chronology, from 1400 ± 10 to $1475\pm?$. Nilakantha was born in 1444, time of death uncertain, possibly after 1530, and Jyeshthadeva was born in 1470 ± 5, date of death again unknown. In line with these dates, we can place Shankara's birth btween say 1510 and 1520 and Achyuta's 20 or 30 years later. With a very slight reservation

in the case of Madhava (see section 4 below), they were all born, did their work, taught their students and prayed to their gods in a cluster of temple villages in the valley of the Nila river.

With the passing of Achyuta, the line broke. He did have disciples, but none who left a mark. The tradition lingered on in fits and starts: a hundred years afterwards, Putumana Somayaji's *Karaṇapaddhati* was written and, another hundred years later, Shankara Varman's *Sadratnamāla*, both of which figure in the title of Whish's 1832 lecture in London. (Whish knew Shankara Varman personally). The locus of the scattered mathematical activity that survived had moved well away from the Nila basin by then but the Nila imprint is as strong as though it was all done in Alattiyur. Shankara Varman, the last of the mathematical Shankaras, in fact did something new that was very much in the Nila mainstream: he refined Madhava's approximations to the π series to the point that he could compute π to 18 decimal places. But times had changed; the Zamorins lost power and prosperity and, after much strife, colonial rule became firmly entrenched.

Thus did the living breath of mathematics finally depart India.

9.3 The Sanskritisation of Kerala

To return to the beginnings, the question now is: what is the connection between what happened in Kerala in the 9th century and in the 14th-15th, between Shankaranarayana and Madhava, or is there a connection at all once we set aside the common allegiance to the Aryabhatan planetary model and Aryabhatan ideology more generally? There is little solid evidence that we can call upon to bridge the gap and what there is is difficult to interpret. Lists of mathematicians and astronomers who might have been from Kerala such as are found in the work of Kunjunni Raja and, especially, K. V. Sarma (cited in the previous section) include a few from the intervening centuries. But they are there primarily on the strength of being mentioned by some Nila authors – particularly Parameshvara and Nilakantha – combined with the negative evidence of their provenance being otherwise unknown. And, of these, only two are mathematicians of some repute: Udayadivakara (11th century) to whom we owe our awareness of Jayadeva as the inventor of *cakravāla* (Chapter 8.3) and Suryadeva (Sūryadeva Yajvan, born 1191). Udayadivakara shares with the 9th century astronomers a strong interest in Bhaskara I and was the author of a commentary on his *Laghubhāskarīya*, the same text that Shankaranarayana had analysed earlier. Suryadeva too wrote on the work of Bhaskara I – so allegience to Bhaskara I might be a connecting thread – but he also chose to go to the original source, in a well regarded commentary on the *Āryabhaṭīya* that Parameshvara made the basis of his own. One of his astrological books was popular in Kerala but that probably says more about Kerala's soft corner for astrology than about his nativity. Mathematical evidence for any decisive influence of either of them on the Nila school is not strong, except on Parameshvara. On the other hand, the

one mathematician who did have a powerful and temporally proximate effect on the Nila work in geometry (see Chapter 8.5), Narayana Pandita, finds no place in the lists, very likely because no Nila text mentions him by name even while building on his work or supplying his theorems with proofs or citing his combinatorial results. This is familiar territory; we are obliged, as in the case of earlier periods, to fall back upon general historical information, from Kerala and outside, for pointers to the currents that converged in the genesis of the Nila school. That history is tied up intimately, once again, with the movement of people bearing with them the traditions of skills and knowledge which first flourished in northern India, to the areas just south of the Vindhya range to begin with and subsequently to the farthest southern reaches of the country (see Chapter 8.1).

The pattern of migration described briefly in Chapter 8.1 began at least as early as Ashoka (3rd century BCE). In keeping with the evangelical tendencies of new faiths, Buddhism and Jainism were well represented among the early migrants to the deep south. But, as we know from Dravidian literature of the first centuries CE (see below), Brahmins were not far behind, in numbers large enough to have had an impact on the ritual and religious practices of the indigenous people. It was not long before the idea was sold to the local chieftains and kings that their legitimacy and worldly power could be given divine backing – an idea that had already been accepted in the north – only through the agency of brahminical rites. The arrangement worked, to the mutual benefit of both parties. In the process, the kings became patrons of brahminical learning as well, just as they had been, in earlier times, patrons of native bards and poets. It is against this background that we have to see the elevation of an astronomer like Shankaranarayana as a protégé of a powerful Hindu king.

Within this overall picture, the specific case of Kerala presents some atypical features, due chiefly to the fact that, about the 8th-9th century, the Brahmin migration to the region grew into a flood that did not abate until the 13th. We have noted earlier that the mathematics of Mahavira did not flourish after him in Karnataka, at least not in the immediate future. (I will argue below (section 4 of this chapter) that Narayana was probably influenced by Mahavira's work but that is 500 years in the future). There were isolated examples, somewhat later, of mathematical scholarship in Andhra and Tamilnadu as well but, again, without any serious follow-up. The situation was quite different in Kerala. Even if we set aside Udayadivakara and Suryadeva, the legacy of the knowledge brought by Govinda (if he was the first to do so) must have lived on through the writings, almost all of it now lost, of obscure but not forgotten authors (see K. V. Sarma, cited above) until it resurfaces, for example in the computation of sines by Parameshvara by means of the half-angle formula in the manner of Bhaskara I (and expressed in the identical *kaṭapayādi* notation of Shankaranarayana). There is also the large number of astrological works of Kerala origin but north Indian inspiration – Varahamihira's above all – produced in the centuries before 1400 (and, also, later). All these astronomers and astrologers have names of classical Sanskrit origin, those of Udayadivakara and

Suryadeva deriving from the sun, like some of their predecessors elsewhere in India.

Some time over these five or six centuries of their increasing domination of the social and cultural life of Kerala, the Brahmins acquired the general apellation of Nampūtiri, often used as part of their names. There are no contemporary records of the particular places in north India they originated from, of what brought them in such large numbers to their new home or of their experiences in first integrating into and then controlling the existing social order. The closest we can get to a narrative in their own words is a longish and very late document in Malayalam, in several versions, prepared probably at the beginning of British colonial rule to establish their notional claim to all the land between the ocean and the mountain chain of the Western Ghats and between Gokarna to the north and Kanya Kumari (Cape Comorin in colonial times) at the southern tip of India. The model for the document may well have been the custom among kings of tracing their lines back to the epic heroes and, as in such panegyrics, much of it is self-serving fiction. But some of the events recounted may have a core of truth in them – since they do not help in advancing their proprietorial claims – such as that they came to Kerala in several waves and had a hard time settling down in an alien landscape overrun by serpents. One historical allusion of interest to us is the tracing of their ancestral clans to Ahichhatra on the Ganga which was always a centre of orthodoxy and learning, especially well known in the 7th century CE as we learn from Huen-Tsang. It ceded scholarly preeminence to other places soon after, among them Valabhi in Gujarat with which it was linked politically. Valabhi also had links with a royal family (the Kadamba, 6th century) whose realm covered the coastal region of Karnataka, with its capital in the hills to the east, and finds a place in some of the epigraphical and literary records from there. Interestingly, it was probably where Bhaskara I taught and it remained a thriving educational centre until the raids of the Islamic kingdom of Sind, established in early 8th century (the Ghazni invasions 250 years later completing the job). This of course proves nothing conclusively; but the presence of astronomer-mathematicians – ardent followers of Bhaskara I in particular – among the migrants to the southwest coast need not cause surprise.

For the first serious studies of Brahmin migration into Kerala, we have to wait till the end of the 19th century, when William Logan, the head of British administration in Calicut, produced his famous *Malabar Manual*. Later work has thrown much light on the dark corners of that history. The following outline, focussed on areas that touch on our concerns, is based on the extensive writings of the eminent historian Elamkulam Kunjan Pillai, done at around the same time as the Tampuran-Ayyar edition of *Yuktibhāṣā*.[10]

[10] Regrettably, he too wrote only in Malayalam. A small selection of his essays was published in English translation a long time back, now out of print. Relying on a variety of primary sources, he put together a coherent model for the political, social and cultural Kerala of the times relevant here and it holds its own even today despite the inevitable revision of many details. For us, he has the added distinction of having recognised the importance

It is not as though the new arrivals found themselves in a land without a cultural identity of its own, a blank slate on which they could inscribe their already ancient ideas on things sacred and secular. The Dravidian civilisation was strong and distinctive and one which valued the finer things in life. From the beginning of the common era, the southern third of India, especially Kerala and Tamilnadu, was home to a vigorous and refined society which cherished its poets and bards as much as it did its warriors and heroes; the early song-poems of this Sangam period are of an astonishing beauty and delicacy, ranging over the concerns of poets everywhere ("Poems of Love and War" as a translated selection is titled). By the 6th century, this golden age began to wind down. The kings and chieftains who ruled on either side of the Western Ghats gave way, on the east coast, to a powerful dynasty with northern affiliations (the Pallavas) and to chaos and confusion on the west, out of which, around the beginning of the 9th century, arose a new imperial dynasty (the Perumals of the second Chera empire) at Mahodayapuram with its sway over most of Kerala, directly or through viceroys and feudatories. It was here that Shankaranarayana found his royal patron and built his observatory. There are no indications, literary or otherwise, that astronomy (or astrology) had a place in the lives of the people during the Sangam period; it was a later import from the north and the fact that astronomer and king could engage in a technically demanding dialogue, in Sanskrit, tells us many things: among others, that this newly arrived science had made much headway by the middle of the 9th century, as had Sanskrit as the language of learning. The Perumals were of indigenous origin but one would not know it from their names or from their favourite deities or from their writings. Culturally, their collective face was turned to the north and the use of Sanskrit was widespread. It cannot be doubted that interest in astronomy was part of a thorough process of Sanskritisation, both in its linguistic meaning and as a misplaced seeking of roots in a Vedic and Hindu antiquity.

How overwhelming the process was becomes clear from what happened over the next few centuries to that primordial cultural artefact that defines a people, their language. Tamil had already begun to make way for a distinctive local dialect and from it, well fed by the rich resources of Sanskrit, emerged a language that is recognisably Malayalam: in the new hybrid all Sanskrit words became legal; even for the names of common objects and for verbs denoting everyday actions, generally resistant to forced change, the Sanskrit alternative was given precedence. Correspondingly, Sanskrit literary forms and norms, including the canons of prosody, were adopted freely. A great deal of experimentation took place and, by the time the great migration drew to its close, certainly by 1400 (that is the time of Madhava), modern Malayalam as a harmonious blend

of Shankaranarayana's *Laghubhāskarīyavivaraṇa* as the first surviving astronomy text from Kerala; he wrote about it and was the moving spirit behind the publication of a critical edition of it. For an excellent capsule-history of the rise and decline of Namputiris, see M. G. S. Narayanan and Kesavan Veluthat, "A History of the Nambudiri Community of Kerala", in [Agni].

of the indigenous language and alien Sanskrit had become the *lingua franca* of Kerala.

But ritual knowledge and Sanskrit were not the only gifts that came from the north, nor language and literature the only spheres that were transformed. The art forms that are today identified uniquely with Kerala were all born out of this encounter. Many among the immigrants were versed in the traditional *śāstra*s: architecture, the martial arts, medical science – though the greater influence here was that of the Buddhists who had arrived earlier – and even magic both 'white' and 'black'. Centres of education (there was one in Tirunavaya), some of them true residential colleges with hundreds of students, were set up in course of time by temples or by munificent notables to promote scholarship in the Vedas and the sciences. There were competitions held in diverse disciplines and purses given to the victors; the court of the Zamorin in Calicut hosted the most celebrated of them, attracting participants from within and outside Kerala. The Zamorin supported an assembly of great scholars known as the eighteen-and-a-half poets – all were skilled versifiers but the half-poet was deficient in the knowledge of Sanskrit. Above all, to be a professor, a *bhaṭṭa* (or Bhattatiri as a title added to the name), became a highly regarded choice of profession for a Namputiri. Kerala of the 14th century was a very different place, with an altogether new personality, from that of the 6th. Seldom in history has the social-cultural matrix that we call a civilisation undergone such a radical transformation, and all without the aid of arms.

The Namputiri dominance of all aspects of public life except the exercise of political power – and even that changed with the passage of time – derived from two circumstances. First of course was their success in pushing the idea that their intercession – their direct line to the divine as it were – was what made the right of kingship legitimate. The other was their sheer numerical strength and that requires an explanation in view of the fact that they formed well under 1 per cent of the population of Malabar in 1881 (when the British conducted the first formal census). It is difficult to be precise about numbers but Elamkulam, whose work I have referred to above, has done a very detailed analysis of the demographic profile of west coast Brahmins over time, with the startling conclusion that Namputiris formed between 20 and 25 percent of the population of Kerala by the 13th century. Of the several factors that go into the analysis, two have a connection to the story of mathematics. The more obvious one is the question of what caused the precipitous fall in the Namputiri population and it has an equally obvious answer. From about 1400 onwards, they followed a particularly regressive form of primogeniture – to safeguard against the fragmentation of family holdings it would seem – in which only the eldest son could marry a Namputiri woman and beget Namputiri children while the younger sons made unofficial but perfectly acceptable liaisons with non-Namputiri women. Their offspring were relegated to a caste lower than Brahmins (but higher than everyone else) and were, in course of time (and if they were male), not excluded from pursuing a life of study. Naturally the relative number of Brahmins dwindled from generation to generation until, by

the 20th century, they were reduced to a tiny minority of well-off landholders (Elamkulam lists several other reasons for the depletion). Among the caste names of the children of unofficial mixed marriages are Vāriyar and Piṣāraṭi; one imagines that if Shankara and Achyuta were born a hundred years earlier, when there were probably enough Brahmin boys to fill the class rooms, they would have languished in their appointed office of service to the temple.

The other factor of relevance arises from the fact that this massive Brahmin migration into Kerala was an extension in time and space of their settlement of coastal Karnataka, the Tulu coast. Considering the route they would have taken, whether from Ahichhatra or Valabhi, that was natural. There are many indications that Brahmin settlements along the west coast covered most of the area later claimed on their behalf, all the way to Gokarna, and that they considered themselves in the beginning as belonging to one community. To a people who did not know the local language, it obviously did not matter much that the northern and southern halves of their new home spoke different languages, Tulu/Kannada and Tamil/Malayalam, and had different historical antecedents. With greater integration into local communities came also linguistic and social differentiation between the Brahmins living in the two regions. In particular, the population dynamics followed different trajectories; Tulu Brahmins have also fallen in relative numbers from their heyday, the reasons for which are reasonably well understood, but nowhere near as dramatically as their Kerala cousins.

For us, more significant than population figures – Elamkulam's primary interest – is the fact that, around the 10th-11th century, the Namputiri identity and name became firmly established as referring exclusively to the Brahmins of Kerala. Tulu Brahmins were never given that name. Families which did cross the divide (the Netravati river just south of modern Mangalore) were given other designations, one of which was Empran or Emprantiri; Emprantiris were not at that time considered inferior to home-bred Brahmins, just a different sort of Brahmins, from a different though closely linked social and geographical milieu. That should set the mathematical-historical bells ringing: one of the few things we know about Madhava is that his family was Emprantiri. We will take a closer look at the use to which we can put this little nugget of information in the next section.

By the beginning of the 12th century, Namputiri dominance of virtually every aspect of life in Kerala became near-total. Temples were built for them in large numbers (among them those of the mathematical villages of the Nila valley) and endowed with lavish land grants for their maintenance which, as happens, became the property of Brahmins associated with them over time. The locals became either their servitors and got assigned to the lowest of the four castes of the north Indian Hindu hierarchy or were declared untouchable outcasts. The penal law was laid down by the Namputiris: one code for themselves and a harsher one for everyone else. At the height of their supremacy, even their regressive marital customs were turned to advantage: younger sons of prominent Namputiri families made alliances with ruling clans, augmenting

their political influence even as their numbers dwindled, and social and family values became corrupted, generation after generation. It used to be said that virtually every raja had a royal mother and a Namputiri father – the throne passed down the female line – and was thus protected both in this world and the other. The quasi-royal lords of Azhvancheri, one of whom was Nilakantha's friend and patron, were for a long time the uncrowned kings of the Namputiris of central Kerala, with a 'palace' close to Alattiyur, never mind the proximity of the secular ruler, the Zamorin. Not for the first time in history nor the last, the highest flights of the mind went hand in hand with degenerate social values.

This was the setting for the second coming of astronomy and mathematics in Kerala and it took place in one of the Namputiri strongholds, the lower valley of the Nila. Some of the 32 villages in Kerala which the Brahmins claimed were given to them as a divine gift in a time beyond memory (another 32 were in the Tulu country) are in the Nila basin. It is a historical fact that two of them, Panniyur and Shukapuram (Chitrabhanu's village) had got into a feud which the then Zamorin settled for them in the 13th century, gaining Brahmin backing and territory at a time when the dynasty was still consolidating its power. The Zamorins had emerged as the most ambitious of the claimants to the remains of the Perumal kingdom after its dissolution in early 12th century. The sands of the Nila at Tirunavaya which earlier hosted a congregation, purpose unknown, every 12 years (named Jupiter's period in Malayalam), were turned into the site for the Zamorin's consecration and the temples taken into his protection, with the blessings of the Brahmins.

There is no doubt at all that the Nila mathematicians did belong to the Nila valley and that they were all Brahmins (or, in its late phase, sons of Brahmin fathers); with the exception of Madhavan Emprantiri, they would have been known locally with 'Namputiri' appended to their names, whatever the ancestral clan they claimed their descent from (the Somayaji in Nilakantha's name just acknowledges his ritual expertise). What might be a reasonable explanation for the sudden flaring of astronomical/mathematical scholarship in the Nila villages in 1350-1400, several centuries after Namputiris had settled these villages in numbers and at least a century of peace after the political events described in the previous paragraph? One would like to believe that the lamp lit by Govindasvami and Shankaranarayana was never fully extinguished, perhaps flaming briefly to life in the work of an Udayadivakara or a Suryadeva. Perhaps Parameshvara, with his strong attachment to Bhaskara I, Govindasvami and Suryadeva came from a family that kept that tradition alive. Whatever the explanation, something happened in the second half of the 14th century to turn the flickering flame into a fire of mathematical creativity; that something came almost certainly from outside Kerala with Madhava as its bearer.

9.4 Who was Madhava? . . . and Narayana?

The defining mathematical event of the five centuries separating Madhava from Shankaranarayana took place far from the geographical limits of Kerala and

that of course was the life and work of Bhaskara II. To keep the chronology in perspective, it is useful to remember that at the time of his birth (1114) the Namputiris were already on the ascendant in Kerala, the Perumals having taken a mortal blow to their empire just a few years before. Suryadeva was born towards the end of the 12th century and there is no Bhaskara II (but plenty of Bhaskara I) in his books, perhaps an indication that they were well separated geographically. Considering how widely Bhaskara's writings came to be disseminated, the greater surprise is that the first commentary on a work of his from (possibly) his natal region or nearabouts (Gaṅgādhara on *Līlāvatī*) is as late as 1420 ([Co], [SiSi-P]); indeed, Parameshvara's, also on *Līlāvatī*, may well be earlier.

Political and military events in the northern Deccan during the period 1200–1350 probably had much to do with the absence of any material related to Bhaskara II datable to this time. As noted in Chapter 8.1, towards the end of the 12th century Muhammad Ghori had devastated all of India north of the Vindhya with a ferocity unmatched since the invasion of the Hunas. Ghori was followed by a succession of sultans belonging to different Afghan-Turkic clans based in Delhi. Two clans, the Khiljis and the Tughlaqs after them, became ambitious and led attacks into peninsular India, plundering and killing on a massive scale. Eventually, the Tughlaqs withdrew to Delhi, leaving behind a fragmented polity in the Deccan, with several Islamic chieftains exercising authority over their individual domains. About the same time (the 1330s), the assertively Hindu Vijayanagar kingdom was founded somewhat further south and quickly grew into a rich and powerful empire and a bastion of Hindu orthodoxy and of Hindu scholarship. Barely 20 years after its founding, what can rightly be called a Vedic workshop was set up, with many pandits working under the direction of the great and famous Sayana; it is to them that we owe the survival of the Vedic corpus as we know it today.

If the pattern of cultural and religious persecution of earlier invasions is anything to go by, traditional scholarship in any field would have found it difficult to sustain itself during the heyday of the Delhi sultans. In the Deccan, the establishment of the Vijayanagar kingdom changed that. The profession of scholarship had by then become a monopoly of Hindu learned men; the Jains had abdicated from their long-held attachment to mathematics, Mahavira being the last of them, and there were hardly any Buddhists left in India, leave alone Buddhist scholars. Practitioners of the astronomical sciences, now part of an exclusively Hindu knowledge system, would have found as warm a welcome from the Vijayanagar dispensation as Sayana and his Vedic experts did. It is reasonable to think that that is what actually happened; that a revival of astronomical-mathematical activity, a coming into the open of what had perhaps remained a private endeavour till then, was made possible under the Vijayanagar umbrella. It was long dead by then in the north but a footnote may be added: astronomy did make a return to the north Indian plains but only towards the end of the 16th century, after the Moghuls had brought a measure of stability and tolerance to the region. The Daivajña family, one of

whom served emperor Jahangir while living in Benares, produced some fine astronomers, especially strong on the work of Bhaskara II. The Daivajña were a community of Brahmins domiciled on the west coast all the way down to the Tulu country (but not Kerala) and among the professions they practised was astrology. It is an intriguing thought that the intimate link with astrology may have been a factor in seeing astronomy over the hard times it went through during this unsettled period; the grandson of Bhaskara II describes himself in the plaque he erected in his ancestor's honour as a *Daivajña*. Even – especially – the powerful liked to know what the stars had in store for them.

In any event, our default choice of provenance for mathematicians who lived between 1300 and 1600, and about whom there is no independent information, is either Kerala or the territory under the control of Vijayanagar – the empire lasted till 1565 – subject of course to revision if and when future evidence comes to light. There is no viable alternative and nothing that we know now contradicts it.

The Vijayanagar empire extended to the west coast, including the Tulu country – the base of the spiritual preceptor of the two founding brothers, Vidyaranya, was on the eastern edge of that region – but not any part of Kerala. The cultural separation between the Brahmins north and south of the river Netravati had become more pronounced with the flow of time but exchanges did take place, more into Kerala than the other way. It is to this diffusion that the relatively small community of the Emprantiris of Malabar can trace its origin. As in other instances of social mobility between castes of roughly equal status, these fine distinctions tended to disappear over time though there are still Brahmin families whose caste apellation remains Emprantiri, as there are other similar subcaste survivals.

This is the background against which we have to assess the reference to Madhava as an Emprantiri. The information appears reliable, at least as reliable as Jyeshthadeva being a disciple of Damodara and the author of *Yuktibhāṣā*, because it comes from the same source, the manuscript giving the partial lineage of the Nila mathematicians (section 2 above): "Mādhavan Ilaññippaḷḷi Emprān" ([Sa-HKSHA]). Ilaññippaḷḷi (or Ilanjippalli) is a village or house name whose identification we will come to in a moment. For the present, let us note that if Madhava or his immediate ancestors came from the Tulu country to settle in the Zamorin's kingdom, the question of the proximate cause of the Nila rejuvenation gets a credible answer: the arrival in its midst of someone who, apart from his own genius, brought with him fresh ideas and fresh knowledge, such as the teachings of Bhaskara II. To remind ourselves again, Parameshvara's commentary on *Līlāvatī*, an unexpected intruder among his dozen or so books on his favourite Bhaskara I, marked the first documented appearance of Bhaskara II on the Kerala scene, maybe the first ever commentary anywhere on any work of his (not counting Narayana's independent monographs on subjects inspired by him). Subsequently, Bhaskara II became an inseparable part of Nila mathematical education. Not only does Shankara devote an

exhaustive *vyākhyā* to *Līlāvatī*, Nilakantha has numerous citations from it and the verse with the list of powers of 10 in *Yuktibhāṣā* is lifted bodily from it.

The fact nevertheless is that, aside from the caste name, there is no other positive evidence for Madhava's Tulu roots. There is, instead, an absence of anything that ties his ancestry to a local background (unlike all his followers), no reference to a favourite temple, for example. No source has anything to say about his teacher(s), and that is very striking in the Nila context where we know the lines of succession in full detail. Nilakantha, otherwise so generous with useful historical and personal asides, is silent on Madhava's intellectual descent: his own is traced to Damodara and Ravi to Parameshvara to Madhava but there it stops. (He even takes the trouble to tell us that Shankaranarayana, 600 years in the past, was a pupil of Govinda). On the other hand, scattered among his rich tributes to Madhava's mathematics in the *Āryabhaṭīyabhāṣya* are a few personal anecdotes which, if we are willing to clutch at straws, may actually point to a connection away from the Nila. In the commentary on *Gaṇita* 12 (the rule for the sine table) he says that he was pleased to have found a new method for the addition formula for sines until he discovered a verse of Madhava's containing the same method; general trigonometric identities are a theme which came to Kerala with Madhava. He also speaks of a meeting he had, while travelling, with an astrologer (the word used is *Daivajña*, not normally employed in Kerala in that sense) born in the Mushika country from whom he learned a stanza composed by Madhava about operations with zero; Mushika is far from the Nila, in the extreme north of Kerala, and shared the Netravati river as the common frontier with the Tulu country. These are at best faintly suggestive of possible northern links, specifically with the mathematics of Bhaskara II (trigonometric identities and operations with zero), that did not exist before Madhava. In the end, however, the only directly personal piece of information we get from Nilakantha is that Madhava was born in a village whose name in its Sanskrit version is Sangamagrama. It is as though Nilakantha was as much in the dark about the pedigree of his great-grand-guru as we are. That is hard to account for if Madhava was from the same milieu as the rest of them with their roots deep in the already ancient Brahmin villages on the Nila.

An appeal to the names of his village and house resolves the mystery partially. Of the two such names we have come across so far, one in Malayalam and the other in Sanskrit, Sangamagrama is not a possible Sanskritisation of Ilanjippalli, obviously not phonetically and also not semantically. A *paḷḷi* originally was a Buddhist retreat house (a *vihāra* in Sanskrit). The word often survived in the names of houses and their compounds after Buddhism's disappearance from Kerala and also lent itself as a suffix to the names of villages – there are thousands of them in southern India – where, it is generally presumed, such institutions once stood. But *ilaññi* is a flowering tree, not the Sanskrit *saṃgama*, a confluence of streams. Here help is at hand from Madhava himself: in his short tract on the moon's positions, *Veṇvāroha*, he says that he, Madhava, was born in a house named (*gṛhanāma*) *bakulādhiṣṭhita . . . vihāra*, which is identical

with Ilanjippalli if we take *paḻḻi* in its meaning of a house, not the whole village (*bakula* is the Sanskrit name of *ilaññi*), i.e., both names refer to his house and Sangamagrama is the village where it was situated.

But where was it? The general view among scholars has always been that Sangamagrama is the town of Irinjalakkuda, some 70 kilometers south of the Nila. It is difficult to see why: there is neither phonetic nor semantic equivalence in the two names; if anything, the similarity, phonetic, is with Ilanjippalli, and even that is not compelling (Malayalam generally does not confuse *r* and *l*). In any case, we have just seen that Madhava's own words contradict the identification unless Ilanjippalli is both a house and its village (that happens occasionally and, astonishingly enough, it happens with the Malayalam equivalent of Sangamagrama, see below) and we discount the significance of the name Sangamagrama altogether. Local historians periodically turn up Brahmin houses in the extended neighbourhood of Irinjalakkuda with names approximating phonetically (but not in meaning) to Ilanjippalli but that does not solve the problem as long as Sangamagrama remains unidentified. House names do change (for example Nilankantha's), village names hardly ever; it is common sense that locating the village must take priority over pinpointing the house.

Tirunavaya itself, right on the Nila, is sometimes called Trimurtisangamam in the temple legends on account of the presence of all three of the main Hindu gods on the two banks, but not Sangamagrama in any record. If we accept, on the evidence of the "Empran", that Madhava's family had its origin in the Tulu country or contiguous parts of the Deccan plateau, another possibility opens up. The area is dotted with villages with *saṃgama* or its Dravidian equivalent *kūḍal* or *kūḍala* in their names (occasionally both as in the most celebrated such place, Kudala-sangama, where Basava, the poet-saint of Karnataka, lived in the 12th century, a contemporary of Bhaskara II) and with a temple consecrated to Saṃgameśvara, the lord of the confluence. Somewhat like Aryabhata the *aśmakīya*, it is entirely reasonable (and a common occurrence in Kerala) for a family to be known by its original domicile. Nilakantha says (in several places) that Madhava was actually born in Sangamagrama, *saṃgamagrāma-ja*. (He also says that Aryabhata was born in the republic of Asmaka, *aśmaka-janapada-jāta*; perhaps *jāta* is only meant as a turn of phrase). Whether Madhava was born in, or in a family from, a village whose name or its Sanskritisation is Sangamagrama, the possibilty that that village was in Karnataka cannot be dismissed.

But, as it happens, there is a much better candidate for Sangamagrama in the Nila heartland itself. This is a village, now a small town, on the river, barely 10 kilometers upstream from Tirunavaya as the crow flies. It is called Kudallur (Kūḍalūr) and its exact literal Sanskrit translation is Sangamagrama: *kūṭal* in Malayalam (as in all Dravidian languages) means a merging or confluence and *ūr* means village. The name comes from its situation at the confluence of the Nila with its main tributary, the river Kunti. If we take Nilakantha at his word, that Madhava was born in Sangamagrama, not merely of it, it would mean

that his family was settled on the Nila before his birth but it really does not make much difference as long as it is accepted that they were recent arrivals who continued in their scholarly vocation. In accordance with custom, Madhava would probably have begun his education within the family. The curriculum would have included, in addition to the standard lessons in Vedic knowledge, language, grammar, etc., the branch(es) of knowledge the family considered its own; if they came from Karnataka, it is not a stretch to suppose that it included the mathematics of Bhaskara II *via* the line of Narayana (see below). Madhava's family tutor, very likely his father, was probably not a highly distinguished scholar, or we would have known about him. On the same ground, if there were astronomers and mathematicians living in the Nila basin at that time, possibly Parameshvara's other teachers, Madhava was not apprenticed to any of them. We should perhaps take Nilakantha's silence as an acknowledgement that Madhava's genius was his own.

To summarise, the most plausible scenario in my view is that the forefathers of Madhava ("Madhavan Empran") migrated from the Tulu coast or thereabouts to settle in Kudallur a generation or two before his birth, in a house whose present identity is unknown. My telling of the story has been convoluted but it is a complicated story, with very few fixed points, and those I have respected. Besides, as in the case of Aryabhata, the magnificence of his mathematical achievements permits and justifies, I think, a certain latitude in our imagined reconstruction of his life.

There are two postscripts to the story of Madhava.

For several centuries, no one seems to know how many, there has existed a Brahmin household known as Kudallur Mana, not in Kudallur village itself nor on the river, but a few kilometers from both. It is a splendid wooded estate now, dominated by a large traditionally built house with many signs still of its antiquity, and it used to be still more extensive in the past. The family history has it that they were originally from Kudallur village and that the main branch was obliged to relocate to the present Kudallur Mana four or five hundred years ago (another example of the persistence through space and time of family origins and family names) for reasons which are not relevant here. What is relevant is that, until about a century back, it was a renowned centre of learning, a *gurukulam* where successive generations of brilliant scholars shared their knowledge and their wisdom with young students and fellow scholars alike.[11] Their speciality was the *vedāṅga*, the six limbs of the Veda, one of which is astral science (*jyotiṣa*). During the time about which there is information, however, their fame rested more on their deep knowledge of the limbs that deal with aspects of language and grammar and of metaphysics. They were champions of the linguistic philosophy of Bhartrhari and literally worshipped

[11] Most of the information given here comes from personal enquiries and contacts; some comes from a fine essay on the Kudallur masters in Alamkod Leelakrishnan's Malayalam book with the (translated) title *Along the Banks of the Nila* (2003), published by D. C. Books, Kottayam.

Patanjali: there is still an antique stone idol of the great grammarian in the house. (The very few images of the conventional gods are recent).

But those who would seek the mathematical spirit of Madhava in this sanctuary of learning or in its earlier setting in Sangamagrama village itself (of the house, nothing remains) need not give up all hope. They may not have been astronomers but the Kudallur Namputiris ("Embrantiri[s] . . . succeeded in later times in securing, or being thought fit to assume, the name of Nambutiri" - Logan in the *Malabar Manual*) were rationalists who believed in and taught the system of metaphysics known as *lokāyata*, the most extremely materialistic and this-worldly of Indian philosophical schools. One of its credos was the contingency of all knowledge: knowledge is not for all time and must be renewed and revalidated constantly through the exercise of our senses and our mind. This is the epistemic current that runs through Nilakantha's last books and it has been suggested that, for all his expertise in ritual and his standing in the Brahmin hierarchy, he can easily be taken for a *lokāyata* sympathiser in his scientific philosophy, at least in his later life. (It has also been said of Aryabhata). The Kudallur masters, while emphatically rejecting any other-worldly influence on the sensible universe, maintained that there was no beginning or end to time and so no need for a supernatural act of creation or annihilation. One would dearly like to know how well they knew their Aryabhata. The house still has a large room called the library, its wooden racks now empty, the manuscript collection dispersed. According to the family, a part of it went to the collection of the old Sanskrit College in Trivandrum which – it is a fact – has several palm leaf manuscripts, including mathematical ones of the Nila school, bearing the Malayalam equivalent of the inscription 'ex libris Kudallur Mana'. Among them is a copy of the *Āryabhaṭīya* dated 1552; Jyeshthadeva was still alive.

It would be a fitting conclusion to the Madhava story if Kudallur was really his home, and to know that he had as brilliant a following in his progeny – in a different field, but that was not unusual – as in his mathematical descendents.

The other remark concerns a more substantive issue and, historically, is more in the nature of a preface than a postscript: who was the conduit for Madhava's familiarity with the works of Bhaskara II? The natural candidate for this role is Narayana (or possibly a school or the family that he belonged to). As we have seen in passing and will discuss in detail below, one interest of Madhava and the Nila mathematicians was the geometry of cyclic figures – to which Bhaskara and Narayana, building on Brahmagupta, made seminal contributions – and its adjunct, trigonometric identities. (The number theory of the quadratic Diophantine equation was not taken up in Kerala, for whatever reason). Narayana was not just the most prominent figure of the period between Bhaskara and Madhava; the broad range of his work and its superlative quality entitle it to a place alongside the very best of Indian mathematics of any period. No personal details are known about him except his father's name, an apparently trivial bit of information but useful all the same (see section 2 of this chapter). He wrote at least two books, *Gaṇitakaumudī* (1356) on general mathematics and a monograph on algebra and number theory, *Bījagaṇitāvataṃsa*

(date unknown). The latter is clearly modelled on Bhaskara's *Bījagaṇita* and the geometry of *Gaṇitakaumudī* (Chapter 8.4) is just as clearly inspired by the geometry in *Līlāvatī*. Narayana's backward linkage to Bhaskara is obvious enough.

So is the forward linkage to the Nila school. To the points already made in Chapter 8.5 (the *Yuktibhāṣā* proofs of his theorems on cyclic quadrilaterals), we can in fact add a new element of an entirely different sort. The last few chapters of *Gaṇitakaumudī* are devoted to a thorough treatment of combinatorial problems. This is somewhat unexpected. Though Varahamihira (in his other avatar as an astrologer-diviner) and Mahavira have some combinatorial results, the subject was not part of mainstream mathematics and had remained more or less the preserve of specialists in theories of prosody, music and dance. In particular, Narayana goes over the recursion relation satisfied by the binomial coefficients ${}^nC_m$ or $N_{n,m}$ used in answering Pingala's question 5 (Chapter 5.5) and relates these coefficients to the so-called sums-of-sums, numbers recursively defined as $S_k(n) = \sum_{i=1}^n S_{k-1}(i)$ (with suitable initial values) of which the lowest nontrivial one, $S_2(n)$, was known to Aryabhata (Chapter 7.1). To anticipate what will be described in detail later, these numbers (for suitable k and n) occur as the exact coefficients in the sine and cosine series before the infinitesimal limit is taken and, in the limit, give rise to the familiar reciprocals of factorials. For the power series, the general $S_k(n)$ is not required, only its value in the limit of large n and that can be obtained by methods already employed in connection with the π series; nevertheless, *Yuktibhāṣā* (as well as Shankara's *Kriyākramakarī*) gives the exact formula. Neither author has an attribution or gives a *yukti*, Jyeshthadeva without explanation and Shankara with a dismissive remark about the "comfort of the not so intelligent". There are more ways than one of deriving the exact expression, all within the demonstrated capability of the Nila mathematicians, but the question remains: where did they get it from? If from Madhava, why is the proof not given? In the case of *Yuktibhāṣā*, the silence is truly puzzling; it is one of only two instances in the whole book in which proof is withheld. The fact is that the only place where the value of $S_k(n)$ is given for general k and n, before *Yuktibhāṣā*, is Narayana's book. It may be surprising that there is no mention of Narayana's name, but no more surprising than that he is not acknowledged in connection with the third diagonal either, where it is absolutely certain that he was the originator of the theorems of which the proofs are found in *Yuktibhāṣā*; indeed, Narayana's name cannot be found in any of the writings of the Nila authors. I tend to the view that they learned about his results not from his books, but from Madhava's mouth, and only selectively (though it still does not explain why no proof is given; we will return to this point later).

From the date (1356) of *Gaṇitakaumudī* we know that Narayana was a generation or two senior to Madhava, perfectly timed to have had a direct impact on him or on his immediate ancestor(s) if the geography allowed it. So, where did he live? Almost certainly not anywhere in the Kerala of his time (as has sometimes been suggested), in view of the fact that Kerala does not know him

by name whether in written records or in orally transmitted folk history. Besides, Pandita was not a common honorific (as distinct from the common noun *paṇḍita* for learned men) among Kerala Brahmins: very occasionally bestowed on great scholars of grammar and philosophy but not on a mathematician or an astronomer. If, on the other hand, he was from Karnataka, it would position him ideally to be both the bearer of the heritage of Bhaskara and its transmitter – as well as of his own mathematics – to Kerala *via* Madhava's family. The idea is not so far fetched as it may seem. Aside from Bhaskara (and, through him, Brahmagupta), the mathematician with whom he has most in common is Mahavira (who himself had a strong commonality of interests with Bhaskara and whose Karnataka connection is not in doubt): Mahavira was one of the very very few mathematicians with an interest in the combinatorics of prosody. In geometry the connection is even stronger, as seen for example in their shared predilection for numerical (integral) geometry; a reading of Sarasvati Amma's book ([SA]), with its systematic ordering of the contributions of individual geometers to each topic, brings this out very clearly. What would be far fetched is to locate him somewhere in northern India, as has also been suggested, for no better reason than that many of his manuscripts were found there. That, as we should know, can be grossly misleading; by the same token, Aryabhata (and lots of others) will qualify overwhelmingly as a son of Kerala (which also, astonishingly, has been suggested in the past). As always in India, manuscripts and the knowledge they contained travelled with their owners, as Bhaskara went to Benares with the Daivajñas and as Narayana might well have. The crucial historical circumstance is that 14th century north India was, after 200 years of the Delhi sultans, an intellectual desert; in contrast, Karnataka was witnessing a reemergence of order during Narayana's very lifetime. With the founding of the Vijayanagar empire came also a cultural and intellectual resurgence driven by people like Vidyaranya, its guide and mentor and a spiritual figure of great authority with his base in an old centre of Hindu scholarship in the hills to the east of the Tulu coast. There is a folk belief that Sayana was Vidyaranya's brother; his Vedic workshop must have been humming with activity during the most productive years of Narayana's life.

As often, a disclaimer is due. We really do not know who Narayana was. Making him an inhabitant of the Vijayanagar kingdom, even if on its periphery, and assigning him the role of the bridge between Bhaskara II and Madhava solve several puzzles but it is nevertheless a speculative and provisional solution, to be confirmed or rejected by future discoveries and analyses.[12]

[12] I have recently learned that there was a well-known Trivikrama Pandita, a religious philosopher, who lived in the 13th-14th century in the extreme north of the present state of Kerala, right on its border with coastal Karnataka. (He even had a son named Narayana Pandita – not, alas, the Narayana of our interest though the two were rough contemporaries). Pandita is one of the common family names of the Daivajña community, to this day. Another small step on the largely untrodden path from Bhaskara II to Madhava?

10

Nila Mathematics – General Survey

10.1 The Primary Source: *Yuktibhāṣā*

The fact that the mathematical writings of the Nila school as a whole exhibit a remarkable unity of purpose does not mean that everything that was not either preparation for or a consequence of the Madhava revolution is absent in them. Collectively, they also serve as a review of the already established body of knowledge; additions and extensions are described, generally in direct continuation of the earlier work but, also, often interestingly fresh. Some of these insights have already been described at appropriate places in this book as the natural culmination of the work of earlier masters. The present chapter is meant to complete the job; it seems best to get everything that is not directly connected with Madhava's programme out of the way before turning to his formulation of calculus on the circle and its supporting mathematics. Perhaps even more than in other mathematical cultures, the conscious clarification and consolidation of already acquired knowledge was always considered in India a precondition for further progress, though parts of it will often be ignored or transcended in the process; in going back to Aryabhatan fundamentals, Madhava bypassed much of what happened in between.

To begin with, a quick recapitulation of some of the topics we have already looked at relating to the amplification and/or reformulation by the Nila texts of material from the *pūrvaśāstra*. Some are foundational principles such as the theorem of the diagonal and its proofs (Chapter 2.2), the decimal place-value principle (Chapter 4.3) and the rules of arithmetic, the rule of three, etc. Some are specific results which had become standard, a good example being the computation of the tabulated sines by means of the half-angle formula which then becomes the starting point in Madhava's interpolation formula. More interesting are ideas and results from the past which are taken up for a deeper

P. P. Divakaran, *The Mathematics of India*, Sources and Studies in the History of Mathematics and Physical Sciences, https://doi.org/10.1007/978-981-13-1774-3_10

treatment as required by the demands of the new problems being addressed, or simply because their proofs were not publicly available. *Yuktibhāṣā*'s more abstract reformulation of the place-value principle in terms of succession (Chapter 4.3) falls in this category as does much of Nilakantha's *bhāṣya* of the *Gaṇita* chapter of the *Āryabhaṭīya*, in particular the geometric demonstrations of the sums of various series (Chapter 7.1) and the derivation of the exact sine and cosine difference formulae (Chapter 7.5). These were essential inputs in the development of calculus and we shall meet them again. On the other hand, the pretty complete account in *Yuktibhāṣā* of the theory of cyclic quadrilaterals as it was handed down by Narayana, including the proofs – available nowhere else – of Brahmagupta's area and diagonal theorems (Chapters 8.4 and 8.5) has no direct relevance in the story of calculus. The rest of this chapter is going to be about some such ideas and results not addressed earlier but which nevertheless form part of the full picture.

The most striking among them is the derivation of the addition formula for the sine which, under the rubric of *jīveparaspara-nyāyam* (an adequate translation will be 'the principle of adjacent chords'), is invariably credited to Madhava. It seems to have had a sensational impact. The derivation is not such a hard problem in itself and the result is not of critical importance for the sine difference formula – which can be obtained without it as Nilakantha demonstrated and Aryabhata probably knew (see Chapter 7.3-4) – and, hence, not also for the derivation of the sine series. Its significance is historical: there is a general belief that it was known to Bhaskara II but no credible indication exists of a proof by him. The possible reason why it was played up – *Yuktibhāṣā* has three different proofs of the result – therefore merits a brief discussion. The other trigonometric topic that will be taken up below, mainly because of its historical (and historiographic) role in the misunderstanding of Madhava's calculus, is his interpolation formula.

Above and beyond such specific results, there are certain very general modes of thought that run through the Nila corpus, chief among them the technique of recursive refining, used in the past for finding exact solutions of arithmetical problems (*kuṭṭaka* and *cakravāla* for example) as well as in iterative schemes of approximation in problems where the exact answer is unknown (and 'unknowable', arithmetically speaking) as in the Bakhshali square root algorithm, or not worth the trouble of working out as in Aryabhata's approximate solution of the second difference equation satisfied by the sine. The recursive mindset has had a pervasive presence in Indian mathematics at all times. In the hands of the Nila mathematicians, often under the general name *saṃskāram*, it became a powerful and versatile formal tool applicable to a wide variety of problems in virtually any mathematical setting. Because its use is not overtly signalled (except in the inductive proofs) its ubiquity, especially in the calculus-related material, is easy to miss; for that reason I have thought it worthwhile to provide a separate section in this chapter outlining the basic premise on which it functions (as well as, in Part IV, a brief excursion into its roots in other areas of early Indian thought). The first recourse to a formal inductive proof (in the

development of the π series) is, in contrast, very conscious and deliberate and it is best to allow its novelty to emerge in that context.

A last general comment has to do with a major conceptual advance in algebra. Madhava's estimates of the error in the π series when it is cut off after an arbitrary finite number of terms needed an extension of the old agebraic considerations (as in Bhaskara's *Bījagaṇita* for example) in a new and much more abstract direction. The error is a function of the cutoff point, an arbitrary positive odd integer, and it is estimated as a rational function of that integer. So, for the first time, the need arose for general definitions of algebraic objects such as polynomials and rational functions in the abstract: the model for a polynomial is a decimally expressed number and the model for a power series is its formal limit as the number is made larger and larger. It is necessary to stress that these are purely algebraic considerations, having nothing to do with calculus *per se*, as we will see more clearly later.

Out of the dozens of books written by those who followed Madhava, the indispensable text for an appreciation of the mathematical culture of the Nila school, in its broad contours as well as in its details, is Jyeshthadeva's *Yuktibhāṣā* (Part I; as before, the restriction to Part I will not be explicitly stated). It will be our guide in the following chapters, supplemented here and there by Nilakantha's *Āryabhaṭīyabhāṣya* for its insights into the Aryabhatan philosophy and, infrequently, by Shankara's writings for material that *Yuktibhāṣā* does not spend much time on. All of them were written a hundred years or more after the passing of Madhava, enough time in which to have absorbed his philosophy and mastered the technical challenges of its implementation.

In its main thrust *Yuktibhāṣā* is a superbly organised monograph on this new mathematics – the first textbook of calculus – beginning with its foundations in arithmetic and geometry and culminating with the sine and cosine series, missing nothing of significance along the way. The innovative ideas that came along with the infinitesimal viewpoint and their mathematical working out are handled with care and pedagogical skill; to cite one example out of many, repeated emphasis is laid on the notion of the limit of a sequence and on the absolute necessity to pass to the limit to get exact answers. (I will provide enough quotations in support of these claims in the following pages). It is a very theoretical book; there are hardly any examples, numerical or not, the exact opposite of mainstream Indian mathematical writing (Aryabhata excepted). The style is business-like, none of the mannerisms of Bhaskara II, no challenges to the intelligent reader and no talking down to the not so intelligent. Every mathematical step is given its *yukti* (except in two instances, interesting for that reason and in themselves, which will occupy us later), and its motivation explained. The general effect is of a master taking his pupils over difficult and unfamiliar mathematical terrain, as in a classroom.

The way the proofs are presented is also different from earlier books, say those of Bhaskara II. First of all, they are complete, not needing the reader or a future commentator to supply the unsaid or unwritten details. Careful attention is paid to structuring them in logically correct sequence, each step building

on what has already been proved: just converting the prose into symbols and equations is often enough to turn them into perfectly acceptable proofs in the modern manner (as we will see in one or two examples). And the reader is left in no doubt where the prefatory explanation ends and the actual proof begins. Despite their differences in style, the closest parallel here is with Nilakantha's treatment of Aryabhata's *Gaṇitapāda*. Perhaps the preference for the word *yukti* over the older *upapatti* in both *Āryabhaṭīyabhāṣya* and *Yuktibhāṣā* marks a recognition of the need for clarity in mathematical methods that Nilakantha so strongly advocated.

Yuktibhāṣā introduces itself with a statement of purpose: to provide an account of "all of the mathematics useful in [the study of] planetary motion according to *Tantrasaṃgraha*". In actual fact it is much more than a compendium of mathematics applicable to astronomy. The first chapter[1] on enumeration and the basic arithmetical operations already gives us a foretaste of the deeper, 'purer' approach we can expect in the rest of the book – see for example the extracts quoted in Chapter 4.2 on the nature of numbers, *saṃkhyāsvarūpam*. The rest of Chapter 1 is more traditional; in particular, it contains a general discussion of the interplay of planar geometry and homogeneous quadratic algebraic identities in two variables, including a proof of the theorem of the diagonal (the one summarised in Figure 2.4(2)), that is a model of clarity and concision. The theme is illustrated in the short Chapter 2 where a series of equations in two variables a and b are solved when any two of the quantities, $a^2 \pm b^2$, ab are given, fairly routine; maybe it is just a token of the continuing interest in the link between the arithmetical/algebraic and geometric aspects of the theorem of the diagonal.

The next two chapters, equally short and unremarkable, deal with the arithmetic of fractions and the rule of three. Chapter 5 on *kuṭṭākāram* is longer and much more interesting in that the mathematics is preceded by a typically careful explanation of the background, namely the astronomical time-keeping problem, illustrating it with numbers (in *kaṭapayādi*) from the real universe, the only place in the whole book where it deviates from its rigid avoidance of illustrative examples. There is a long extract from *Līlāvatī* on the solution of the linear Diophantine equation but the explanation following is slightly biassed towards expressing it in terms of continued fractions (see the Appendix (in English) in [YB-TA]). How well-versed the Nila mathematicians were with the theory of continued fractions will turn out to be an important question when we come to the truncation corrections in the π series. The texts unfortunately do not have much to say about it but *Yuktibhāṣā*'s familiarity with the connection

[1]The references here and in the rest of the book to chapter and section numbers are to the text as in Sarma's book [YB-S]. The organisation of sections and subsections in the Tampuran-Ayyar edition [YB-TA] is not always the same; besides, it does not number them. Translations unless otherwise indicated are my own. The reason why Sarma gave the manuscript the title *Gaṇita-Yukti-Bhāṣā* remains mysterious. In the Nila tradition and in writings from Kerala more generally, whether in Sanskrit or Malayalam, it has always been known as *Yuktibhāṣā*.

between the Euclidean algorithm and continued fractions is an indication that the subject was not a closed book.

Chapters 6 and 7, the last two, form the core of the book. They are by far the longest, taking up between them about two-thirds of the whole work. All of Madhava's new mathematics is to found in these chapters and the only thing in them that is not directly connected to it is the theory of cyclic quadrilaterals in the second half of Chapter 7. Chapter 6 titled "Circumference and Diameter" is concerned with the π series, its generalisation to the arc-tangent series for arbitrary angle and the problem of improving its numerical efficiency by rearranging the series or by estimating the error when the summation is stopped after an arbitrary finite number of terms. The infinitesimal philosophy makes its first ever explicit appearance in a text here – keeping aside Aryabhata's elliptic *yatheṣṭāni* – and, accordingly, is given a meticulously careful explanation every time it is invoked. But, before that, there is the inevitable recollection of the diagonal theorem with yet another proof (the one illustrated by Figure 2.3), followed by the method of computing the circumference as the perimeter of regular polygons of successively doubled sides starting with the circumscribing square; it is brought to a close with the sentence "[I] now relate a method of producing the circumference from any diameter, without taking square roots". Thus was the birth of calculus first announced publicly to the world.

Chapter 7 is divided roughly equally between the mathematics of the sine and cosine series and the geometry of cyclic quadrilaterals, in that order, with the area and volume formulae for the sphere bringing up the rear. The proofs of Brahmagupta's theorems that I have described in Chapter 8.5 – and the attendant geometry of Narayana's third diagonal – dominate the second half of the chapter. The subject is introduced with the statement that the properties of cyclic quadrilaterals offer a way of determining the chord without reference to the radius – the Euclidean (or Ptolemaic) viewpoint, one might say – which is probably meant to set it apart from Aryabhatan trigonometry; in other words, the focus is on the geometric properties of circles and their chords without introducing the half-chord, without trigonometric functions. The more specific message seems to be that such results are independent of the units used to measure the arc, unlike the series which hold only when angles and chords are measured in the same units (hence Rsines). They have no direct role in the infinitesimal trigonometry and, after having done all the work, *Yuktibhāṣā* puts them to no use, except in one proof of the addition theorem of sines. But for that proof, and after the account given in Chapter 8.5, there will be no need to return to the subject.

As regards the first half of the chapter, it begins with a perceptive introduction to general trigonometry, paying special attention to the definition of the sine and the cosine for arbitrary angles – i.e., in all four quadrants – highlighting thereby their periodicity and complementarity properties. The half-angle formula is derived and the sine table, obviously meant to be prepared using that formula, gets a passing mention – "since they were studied in earlier texts" referring, perhaps, to Bhaskara I – followed by the first order

interpolation formula for the preparation of a finer table at angles which are half the canonical angles. What follows next is the mathematics of the sine and cosine series described, as in the study of the π series, with great care. It is made clear at several points that the ultimate goal is to find exact expressions for $\sin\theta$ and $\cos\theta$ for any angle θ, i. e., to determine them as functions of θ just as, in the arctangent series, the aim is to determine θ as a function of $\tan\theta$ or $\cot\theta$ (Chapter 11.5 below) in all four quadrants. But, unlike in the latter problem, the solution cannot be reduced to a simple quadrature. The alternative strategy adopted is the discrete equivalent of solving the differential equation satisfied by the function, combined with a delicate handling of the necessary limiting operations. Aside from these fine points of theory, Chapter 7 consequently contains some of the most technically demanding mathematics in the whole book, indeed in all of Indian mathematics. That, in the end, we are able to reconstruct the entire chain of arguments is a tribute to both the clarity of thought and the explanatory skills of Jyeshthadeva.

The chapter (and the book) ends with derivations of the formulae for the surface area and volume of the sphere. There are no series here and that is of great value since it allows an assessment of whether Madhava's ideas and methods really qualify as calculus, without other distracting novelties such as infinite series coming in the way. They are in fact perfect illustrations of why calculus was invented: to deal with problems involving curvature (*Yuktibhāṣā* says so). The infinitesimalisation is, once again, carefully explained and carried out; Bhaskara II is not mentioned but it becomes clear along the way why he could not back his correct formulae with proper proofs (see Chapter 7.5) and what had to be done to remedy the situation. *Yuktibhāṣā*'s approach is identical in all essentials with how it is done in today's classrooms, with some interesting deviations in the details, as we shall see. The differential equation satisfied by the sine again plays a role here but, as far as I know, no one has attributed these particular derivations to Madhava. No matter; they are the best counterarguments to those who have questioned the calculus credentials of Nila mathematics.

In the end, one cannot help noticing that much of the contents of Chapters 6 and 7 is not essential for astronomy, despite *Yuktibhāṣā*'s opening sentence about "mathematics of use in the study of planetary motion". For the level of accuracy demanded by observations – no instruments other than line-of-sight mechanical devices – a sine table accurate to the minute was adequate and for that accuracy, there was no need for a value of π beyond the five significant figures of Aryabhata. The need for a table more closely spaced than 225 minutes, a large angle even for the naked eye, must have been felt and that was taken care of by the interpolation formula. Neither the π series nor the sine series would have been of any use to the practising astronomer. Calculus was not born out of a pressing practical need; from all we know, Madhava was barely an astronomer.

And that brings up a final question: how faithful a picture do we get of Madhava's thought processes when seen through the lens of Jyeshthadeva's

account of the results they led to? It is here, in assessing Jyeshthadeva's own contributions in giving these results a solid theoretical foundation and in fitting them into a unified and coherent mathematical structure, that we are seriously handicapped by the lack of anything mathematical in Madhava's own words. It is not to be doubted that the individual results ascribed to Madhava were his own but that is all they are, a collection of results, brilliant in themselves. From Nilakantha's *Āryabhaṭīyabhāṣya* one comes away with the impression that Madhava's preferred way of communication was to encode his discoveries in individual aphoristic verses – many, perhaps even most, of which are quoted in *Yuktibhāṣā* and/or *Yuktidīpikā* – which he then scattered about. Perhaps he himself never wrote a complete account of all his original work. Certainly his only disciple Parameshvara did not. It is tempting to think that the structural framework in which *Yuktibhāṣā* holds it all together was at least partly Jyeshthadeva's creation and that it reflects the epistemic concerns – the imperative of logical clarity and demonstrability and the transformation of informal *upapatti*s into formally impeccable *yukti*s – of his guru Nilakantha in his late years. To a modern reader willing to make the effort to penetrate the narrative style of *Yuktibhāṣā* and turn it into symbols and equations, as striking as its contents is the sense of modernity it conveys.

10.2 Geometry and Trigonometry; Addition Theorems

This section is mostly about the addition theorem of sines:

$$\sin(\theta \pm \phi) = \sin\theta\cos\phi \pm \cos\theta\sin\phi,$$

one of its proofs and, to a limited extent, its history in the Indian context. It covers the cosine theorem

$$\cos(\theta \pm \phi) = \cos\theta\cos\phi \mp \sin\theta\sin\phi$$

also tacitly, though it is not always separately mentioned. As we saw in connection with Playfair's justification of Aryabhata's rule (Chapter 7.4), a knowledge of it was not necessary for working out the rule, the difference formula was enough. But it seems to have been recognised progressively that the general addition formula could be used to derive other useful trigonometric identities as special cases or through algebraic manipulation, the sine difference formula being the first instance. I will argue here that the theorem, very probably, did not get a proof, possibly not even a clear statement, before Madhava. Going by the enthusiasm with which it was hailed, one would guess that something along its lines was anticipated; and once guessed, it would have been easy to verify it from an inspection of the sine table. The statement of the theorem occurs in Nilakantha's *Tantrasaṃgraha*, attributed to Madhava under what became its formulaic name *jīveparaspara nyāyam*, but *Yuktibhāṣā* is apparently the first

text to prove it and it may even be that it was not explicitly mentioned in the pre-Nila literature (see comments at the end of this section) precisely because there was no proof. Its historical interest really derives from its late arrival on the scene; everything that it was used for had already been done by other methods.

But, before that, I describe two other illustrative results from general circle geometry ('general' meaning, as earlier, not tied to the limit of small arcs or angles), one due to Parameshvara and the other from *Yuktibhāṣā*. The idea of transpositions in cyclic quadrilaterals figures in both of them. In fact the first, whose proof involves the third diagonal explicitly, is a corollary of Narayana's theorem of three diagonals and is hence the earliest documented evidence for Narayana's direct influence on Nila mathematics – passed on by Madhava to his disciple? The second result occurs as a lemma in *Yuktibhāṣā*'s proof of Brahmagupta's area theorem – and, to that extent, is also a token of Narayana's influence – and we have already met it in that guise in Chapter 8.5. The reason for returning to it is that it translates immediately into a trigonometric identity, perhaps the simplest example of alternative geometric and trigonometric formulations of theorems on cyclic figures.

Parameshvara's result is an expression for the circumdiameter as a function of the sides alone. At first look it is a trivial corollary of Narayana's theorem of three diagonals, $4\mathcal{A}\mathcal{D} = d_1 d_2 d_3$ in the notation of Chapter 8.5. Substitute for $\mathcal{A}$ and d_1, d_2, d_3 their expressions in terms of the sides (Brahmagupta, generalised to the third diagonal, Chapter 8.5) and we get

$$\mathcal{D}^2 = \frac{(a_1a_2 + a_3a_4)(a_1a_3 + a_2a_4)(a_1a_4 + a_2a_3)}{4(s - a_1)(s - a_2)(s - a_3)(s - a_4)},$$

which is Parameshvara's formula. We have to remember however that the three-diagonal theorem is also an input in the computation of the individual diagonals which Narayana does not have – it is not clear whether it was known to Parameshvara – and so it is not such a trivial consequence of what is actually in Narayana. Anyhow, it is a beautiful formula, pleasingly symmetric in the sides, reflecting the underlying transposition invariance and fully in line with Narayana's own work.

The geometric version of the second example is elementary and we have already used it in the evaluation of the parameter XY that measures the deviation of a cyclic quadrilateral from being Brahmagupta (Chapter 8.5). Referring to Figure 8.6, it says simply that if BX is an altitude of triangle ABC (with $AB > BC$), then $AB^2 - BC^2 = AC(AX - CX) = (AX + CX)(AX - CX)$. It becomes interesting if we think of AC as one diagonal of the cyclic quadrilateral $ABCD$ of Figure 8.6 and consider the symmetric trapezium $AB'BC$ where B' is the new vertex of the transposed quadrilateral as in the figure, DB' being the third diagonal. Imagine drawn the altitude $B'X'$ of the triangle $AB'C$. Then $AX - CX = AX - AX' = XX' = BB'$, which is the full chord of $\text{arc}BB' = \text{arc}AB - \text{arc}AB' = \text{arc}AB - \text{arc}BC$. Since the relation $AB^2 - BC^2 = (AX + CX)(AX - CX)$ is valid for any three points A, B, C on

the circle with BX perpendicular to AC, we can reexpress it in terms of arcs and chords as

$$\gamma(\alpha_1)^2 - \gamma(\alpha_2)^2 = \gamma(\alpha_1 + \alpha_2)\gamma(\alpha_1 - \alpha_2)$$

for arbitrary arcs α_i, where $\gamma(\alpha)$ is the full chord of α. A redefinition leads to the equivalent identity

$$\gamma(\alpha_1)\gamma(\alpha_2) = \gamma((\alpha_1 + \alpha_2)/2)^2 - \gamma((\alpha_1 - \alpha_2)/2)^2,$$

which is how *Yuktibhāṣā* states the proposition: "The product of two arbitrary chords is the difference of the squares of chords associated with half the sum and [half the] difference of the corresponding arcs."

The point of *Yuktibhāṣā*'s reformulation of a geometric property of triangles as a property of arcs of the circumscribing circle was first noted in modern times by Tampuran and Ayyar ([YB-TA]). It turns circle geometry into trigonometry: we only have to rewrite it in terms of half-chords. If the arc α is identified with the angle 2θ it subtends at the centre, we have $\gamma(\alpha) = 2\sin\theta$ for a circle of unit radius and the *Yuktibhāṣā* proposition becomes

$$\sin^2\theta_1 - \sin^2\theta_2 = \sin(\theta_1 + \theta_2)\sin(\theta_1 - \theta_2)$$

or, equivalently,

$$\sin\theta_1 \sin\theta_2 = \sin^2((\theta_1 + \theta_2)/2) - \sin^2((\theta_1 - \theta_2)/2).$$

A first remark is that the result can be easily derived, as is done in school textbooks, from the addition theorem for the sine (together with the complementarity of the sine and the cosine); presumably it predates Madhava's proof of the addition theorem. Jyeshthadeva's proof of Brahmagupta's formula for the area, which comes after this proposition in the book, uses only the elementary geometric version of it.

A second remark touches on a more fundamental issue, the relationship between cyclic geometry and trigonometry, that came up in our earlier discussion of what constitutes (or should constitute, logically and historically) the discipline of trigonometry (see Chapter 7.3 and Chapter 7.5). A cyclic quadrilateral of a given radius is defined (up to ordering of sides) by three sides or the three angles they subtend at the centre, the two being related by Aryabhata's half-chord idea: $a_i = 2\sin(\theta_i/2)$ (for unit radius). Any quantity of interest, lengths of diagonals, area, etc., expressible as a function of the sides alone can in principle also be expressed as a function of angles alone (and *vice versa*) and any relation among the sides can be rewritten as a trigonometric identity involving three angles, the fourth angle being $\theta_4 = -\theta_1 - \theta_2 - \theta_3 \pmod{2\pi}$. The *Yuktibhāṣā* identity we have just looked at comes from this correspondence applied to a cyclic quadrilateral with extra symmetry, namely a symmetric trapezium ($AB'BC$ of Figure 8.6). There are only two independent sides and the resulting trigonometric identity involves just two angles.

Another illustration of this equivalence is provided by one of *Yuktibhāṣā*'s proofs of the addition theorem itself. These proofs are worked out for the case

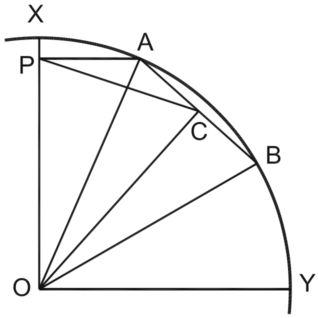

Figure 10.1: Addition theorem from Ptolemy's theorem.

of θ and ϕ being multiples of the standard angular step $\pi/48$ but it is clear that they are meant to hold, and are valid, for arbitrary angles. (This is not the only time that Jyeshthadeva worked with the standard step in place of a general angle; he explains elsewhere that it is done for the sake of definiteness). Two of the proofs are based on ingeniously constructed similar triangles, reminiscent of Nilakantha's proof of the sine difference formula (Chapter 7.4, Figure 7.4) but a little more intricate. (The interested reader will find them in modern notation in the Explanatory Notes of [YB-S]). The third proof, given below for the sine of the sum, is far more interesting; it establishes that the addition theorem is a consequence of the product formula for the diagonals of a cyclic quadrilateral (Ptolemy's theorem) and conversely; yet another illustration of the mutability of cyclic geometry and trigonometry.

In Figure 10.1, the arc XY is a quadrant of a circle of unit radius centred at O. Measuring angles from X, A and B are points corresponding to angles θ and $\theta + 2\phi$ respectively, OC bisects AB (perpendicularly) and AP is perpendicular to OX. Then $AP = \sin\theta$, $AC = \sin\phi$, $OP = \cos\theta$ and $OC = \cos\phi$. Moreover, $OPAC$ is a (pre-Brahmagupta) cyclic quadrilateral of which the diagonal $OA = 1$ is a circumdiameter. From Ptolemy's theorem then, $CP = OA \cdot CP = AP \cdot OC + OP \cdot AC = \sin\theta\cos\phi + \cos\theta\sin\phi$. We only have to show now that $CP = \sin(\theta + \phi)$. *Yuktibhāṣā* indicates that it can be done by a method it has already used involving similar triangles but Tampuran and Ayyar in [YB-TA] point out a simple and elegant argument, very much in line with Aryabhatan trigonometry, depending on the scale-invariance of the sine or, equivalently, the scaling of Rsine (more generally, chords) with the radius. To spell out the details, PAC is an arc of the circumscribing circle (not drawn in the figure) of $OPAC$, of radius $1/2$, subtending an angle $\theta + \phi$ at the point O on it. So it will subtend an angle $2(\theta + \phi)$ at its centre (the midpoint of OA); hence its full chord is given by $CP = 2 \times (1/2) \times \sin(\theta + \phi)$ where the factor $1/2$ is the radius of the relevant circle.

Yuktibhāṣā does not say that this proof is due to Madhava; and *Tantrasaṃgraha*, which does say that the theorem is due to Madhava, has no proofs. If Madhava did devise this particular proof, it would mean that he was

thoroughly familiar with Brahmagupta's theorems, presumably *via* Bhaskara II and Narayana.

The addition theorems also have a post-Madhava history and that history has an intriguing pre-Madhava connection. In the 17th century, Munishvara (Munīśvara) wrote a commentary named *Marīci* on the *Golādhyāya* part of Bhaskara II's *Siddhāntaśiromaṇi* which, according to him, has a chapter titled *Jyotpatti*, "Origin (or production) of sines", on general trigonometry, containing among other things a statement of the addition theorem of sines, without proof.[2] The commentary, however, has a remarkably unusual attempted proof, starting with the fact that $(\pm\sin\theta, \pm\cos\theta)$ for arbitrary θ are solutions of $x^2+y^2=1$. The last is formally the quadratic indeterminate equation of Brahmagupta for $N=-1$: $-x^2+1=y^2$. It has only the trivial solutions $(\pm 1, 0)$ and $(0, \pm 1)$ in integers (corresponding to $\theta =$ an integral multiple of $\pi/2$) but, as we noted in Chapter 8.2, the principle of *bhāvanā* holds for real solutions as well. Composing the solutions $(\pm\sin\theta, \pm\cos\theta)$ and $(\pm\sin\phi, \pm\cos\phi)$ pairwise among themselves, we see that $\pm(\sin\theta\cos\phi \pm \cos\theta\sin\phi)$ and $\pm(\cos\theta\cos\phi \pm \sin\theta\sin\phi)$ are all solutions of the same equation, i.e., that they are all sines or cosines of some angle χ. Munishvara asserts that χ is the appropriate sum or difference of θ and ϕ but there is no proof of it; it would seem that he was trying to supply a proof for a result that he knew was true. (Bear in mind that he lived about a century and a half after Madhava). It is an interesting observation nevertheless that *bhāvanā* applied to real numbers may have a use in the study of periodic functions; that, for example, additive *bhāvanā* (for $N=-1, C=1$) on $(\sin\theta, \cos\theta)$ is just addition of angles: $(\theta,\phi) \to \theta+\phi$, once the addition theorem is known.

The really interesting question is whether Bhaskara himself had a proof of the addition theorem; familiar as he was with the sine table, he may well have guessed its correctness as he did the sphere formulae. About a proof, however, there are several reasons for scepticism, all boiling down to the authorship of the *Jyotpatti*. It is purported to be the last chapter of the part of the *Siddhāntaśiromaṇi* dealing with the sphere and, as was noted by GuptaGuptas (citing Sudhakara Dvivedi), it comes after the end-marker of the text, designated as such by the author himself. Most of its contents is a hodge-podge, out of character with everything we know and admire about Bhaskara and his style; for example, sines are to be determined empirically, by actually drawing a circle and its chords, as though the author was ignorant of the half-angle formula and Aryabhata's rule (but not the cosines which are to be computed from the sines). There is a miscellaneous collection of trigonometric odds and ends, without even an indication of the proof for any of them (they include the remarkable result $\sin\pi/5 = \sqrt{(5-\sqrt{5})/8}$; paradoxically, whoever was the author had access to trigonometric results going well beyond the half-angle formula) and that is also uncharacteristic. Elsewhere in *Siddhāntaśiromaṇi* Bhaskara

[2] R. C. Gupta's article, "The *Marīci* Commentary on the *Jyotpatti*", Indian J. Hist. Sci., vol. **15**(4) (1980), p.44 and references therein form the basis for the following remarks.

says that astronomy not supported by mathematical demonstration will get no respect from the learned, it is "like food without butter"; *Jyotpatti* itself has a prefatory line about how its mastery is the hallmark of a great teacher, but the text belies that claim.

But a much more persuasive reason to doubt Bhaskara's authorship of *Jyotpatti* and, hence, of the addition formula, is that he is not mentioned in connection with them in any of the Nila texts. They are otherwise rich in references to him and his books – by far the most profusely quoted of all the earlier masters – but there is none to any result occurring only in *Jyotpatti*. Whenever there is an explicit attribution of *jiveparasparam*, it is invariably to Madhava, for example Nilakantha in *Tantrasaṃgraha* (*mādhavoditam*, "stated by Madhava") and in *Āryabhaṭīyabhāṣya* (*mādhavanirmitaṃ padyam*, "a verse composed by Madhava"). It seems far-fetched to suppose that Nilakantha, especially, would have been ignorant of the fact if Bhaskara had already stated, if not proved, these identities.

In the big scheme of things that Madhava set in motion, the addition formula are a small detail. The *Jyotpatti* episode is interesting for the light it throws on the ambiguities inherent in the reading of isolated texts by relatively minor authors. It is a very short work and is not mentioned by anyone (except purportedly by Bhaskara himself in his own commentary on *Siddhāntaśiromaṇi*) before the 17th century and nothing is known about where Munishvara might have got it from. We have only his word that it was authored by the universally admired Bhaskara. The circumstantial evidence does not support his word; the Daivajña family, the custodians of Bhaskara's work and living in Benares, does not mention it (as far as I know). We have also to keep in mind that Munishvara lived long enough after the decline of the Nila school for bits and pieces of their discoveries, though not a coherent account apparently, to have diffused to northern India – we must not forget the astrolger (*Daivajña*) who repeated one of Madhava's mathematical verses to Nilakantha at the northern extremity of Kerala – where, under the Mughals, a revived mathematical life had just taken hold. The abstract-geometric viewpoint of the Nila school, as evidenced by the *Yuktibhāṣā* proofs, would have been completely out of tune with his own which was more utilitarian and computational. In any case, the *Marīci* commentary is a confused work; among other things, it explains that the reason the cosine formulae are not found in *Jyotpatti* is that cosines can always be computed from the sines which are, we have already been told, to be determined by drawing figures. To use his own words, he probably found trigonometry *atidurga*, extremely hard of access.

10.3 The Sine Table; Interpolation

Even if we disregard the extreme detail in which Chapter 7 of *Yuktibhāṣā* deals with "the theory of sines" (*jyānayanam*), its treatment of that old staple, the table of sines, is brief and unexciting, as though it wants to be quickly rid of

a topic which does not hold much interest for it. It occurs at the very beginning of the chapter with, first, a passing acknowledgement of the calculation of $\sin m\epsilon, \epsilon = \pi/48$, by means of the half-angle formula and complementarity, both of which are stated explicitly, maybe because they are fundamental properties. The values of the sines of $\pi/6$ and $\pi/4$ are given but nothing is said about how to use them, only that "some sines can also be produced this way". More informative are some of the casual remarks that come along with the mathematics: "In computations involving planets, *only the half-chord is of use* (my italics); that is why the *arddhajyā* (the extra *d* reflects a feature of Sanskrit words taken over by Malayalam) is called *jyā*"; or, "When we imagine a circle drawn on the ground . . . ". One has the impression of listening in on a class in progress.

This is followed by an account of a somewhat unexpected method for preparing the table using the difference formulae. Since the difference of two consecutive tabular sines, $\sin(m+1)\epsilon - \sin m\epsilon$ is proportional to the cosine at the midpoint, $\cos(m+1/2)\epsilon$, start with $\sin(\epsilon/2) = \epsilon/2$ and $\cos(\epsilon/2) = \sqrt{1-\epsilon^2/4}$. Then compute $\sin\epsilon = 2\sin(\epsilon/2)\cos(\epsilon/2)$ from which we know $\cos\epsilon$ as well. Now apply the sine difference formula to the half steps: $\sin(3\epsilon/2) - \sin(\epsilon/2) = 2\cos\epsilon\sin(\epsilon/2)$, to get the value of $\sin(3\epsilon/2) = 2\cos\epsilon\sin(\epsilon/2) + \sin(\epsilon/2)$ (and similarly for the cosine) and so on. A couple of remarks can be made. The formulae for the differences are not assumed to be already known; they are worked out geometrically for the half steps following Nilakantha's general method (Chapter 7.4, Figure 7.4). Because the midpoints of the canonical division are involved, the final table is twice as fine as the conventional one; effectively, the step size has become $\pi/96$. Jyeshthadeva could well have started with the approximation $\sin\epsilon = \epsilon$ and worked with the two-step difference if he just wanted the standard table; clearly, he knew that the method was independent of the step size and that the table could be made for as fine a division of the quadrant as desired. He says elsewhere (in connection with the sine series) that the step size is not fixed, the choice of ϵ being just a token of its familiarity. This may appear to be an obvious point, but, equally obviously, a knowledge of differences for arbitrary (arbitrarily small) angles is critical for a calculus of trigonometric functions. It is the kind of insight missing from Bhaskara II's writings on the subject.

But, curiously, second differences and their self-proportionality property, going back to Aryabhata and so effectively employed in the calculus to come later in the chapter, do not make an appearance here. What we have instead is a reminder of the necessity of resorting to finer and finer subdivisions of the arc to handle curvature: "Since the [effect of the] curvature is small [for an arc of $\pi/96$], we can suppose that the arc is equal to the chord". The general warning about the danger of using the rule of three when dealing with curves that I have already quoted at the end of our Chapter 7.5 comes just before and is thus reinforced.

Knowing the sines and cosines of $m\pi/(2p), 1 \leq m \leq 2p, p$ even, *Yuktibhāṣā* next describes, even more briefly, an interpolation formula due to Madhava for approximating their values at an arbitrary intermediate point $m\pi/(2p) + \delta, 0 < \delta < \pi/2p$. It is done for the canonical step $p = 24$ but works equally well for

any p. The input is again the difference formula and it is, once again, implicitly derived in this new setting while working out the first order (linear in δ) of interpolation. We can simplify the argument by omitting the geometry and applying the difference formula directly to $\sin(m\epsilon + \delta) - \sin m\epsilon = \sin(m\epsilon + \delta/2 + \delta/2) - \sin(m\epsilon + \delta/2 - \delta/2)$ and setting $\sin(\delta/2) = \delta/2$ to conclude

$$\sin(m\epsilon + \delta) = \sin m\epsilon + \delta \cos(m\epsilon + \delta/2)$$

and, similarly,

$$\cos(m\epsilon + \delta) = \cos m\epsilon - \delta \sin(m\epsilon + \delta/2).$$

The first order (linear) interpolation formula results from dropping $\delta/2$ in the argument of the second term on the right:

$$\sin(m\epsilon + \delta)_1 = \sin m\epsilon + \delta \cos m\epsilon,$$

$$\cos(m\epsilon + \delta)_1 = \cos m\epsilon - \delta \sin m\epsilon.$$

The approximation improves with increasing m but obviously is not very good in general, even if p is taken to be large. The equally obvious remedy is to work to the second order (and higher orders if the need is felt) in δ.

The method of extending the interpolation to higher powers of δ is of great methodological interest, quite apart from its practical usefulness. The exact equation for $\sin(m\epsilon + \delta)$, reducing its evaluation to that of $\cos(m\epsilon + \delta/2)$ (and similarly for the cosine) is an open invitation for the application of the general approach that we have called recursive refining (*saṃskāram* in the Nila school's preferred terminology). The implementation of the method in this particular example is an excellent case study of its strengths as well as of the misunderstandings that it sometimes gives rise to, making it worth our while to spend a little time on it. *Yuktibhāṣā* instructs us to treat $\sin(m\epsilon + \delta/2)$ (and similarly the cosine) as the new 'desired' but unknown quantity and apply the difference formula to it, i.e., replace $\cos(m\epsilon + \delta/2)$ by $\cos m\epsilon - (\delta/2)\sin(m\epsilon + \delta/4)$ and drop $\delta/4$ from the argument of the sine, resulting in the second order interpolation

$$\sin(m\epsilon + \delta)_2 = \sin m\epsilon + \delta \left(\cos m\epsilon - \frac{\delta}{2} \sin m\epsilon\right).$$

At this point Jyeshthadeva offers the advice: "If even this (the resulting precision in the sines) is not enough, halve the chord again", closing the subject abruptly with a verse from the *Tantrasaṃgraha*. His interest in the procedure is stricly utilitarian, arising from the fact that $\sin m\epsilon$ is known from the table and δ is small enough for $\sin(\delta/2) = \delta/2$ to be a good approximation.

The approximate interpolation can be repeated as many times as desired, keeping in mind that its accuracy is really controlled by the approximation in the first step, that of replacing $\sin(\delta/2)$ by $\delta/2$, no matter how many iterations

are subsequently made. But the underlying principle is itself independent of such practical details. The exact difference formulae

$$\sin(\theta+\delta) = \sin\theta + 2\sin\frac{\delta}{2}\cos\left(\theta+\frac{\delta}{2}\right)$$

and

$$\cos(\theta+\delta) = \cos\theta - 2\sin\frac{\delta}{2}\sin\left(\theta+\frac{\delta}{2}\right),$$

valid for general angles θ and δ, are simultaneosly 'solved' formally by a process of repeated cross-refining: in the language employed in the iterated refining of square roots in the Bakhshali manuscript (Chapter 6.5) and possibly in the *Śulbasūtra* (Chapter 2.5), $\sin\theta$ is the first guess in the evaluation of $\sin(\theta+\delta)$, $2\sin(\delta/2)\cos(\theta+\delta/2)$ is the exact difference (replacing the first factor by δ is an option available when δ is small), dropping $\delta/2$ in the argument of the cosine leads to the first correction, and so on. It is in computing the second correction that the difference formula for the cosine has to be invoked, thereby turning the procedure into the simultaneous evaluation of both the sine and the cosine; but for this added subtlety, the refining of trigonometric functions is no different in philosophy from the refining of the square root function.

The distinction between the sequence of exact relations underpinning the general method of *saṃskāram* and its use in approximation schemes is not always appreciated. In the problem at hand, the third step for example gives the identity

$$\sin(\theta+\delta) = \sin\theta + 2\sin\frac{\delta}{2}\left(\cos\theta - 2\sin\frac{\delta}{4}\sin\left(\theta+\frac{\delta}{4}\right)\right)$$

and so on; after k steps, we will have an exact expression for $\sin(\theta+\delta)$ involving the unknown sin or cos (depending on the parity of k) of $(\theta+\delta/2^k)$. But it will also involve the products $\sin(\delta/2)$, $\sin(\delta/2)\sin(\delta/2^2), \cdots, \sin(\delta/2)\sin(\delta/2^2)\cdots\sin(\delta/2^i)$, for $i \leq k$. However many times iterated, it is the supposed smallness of δ, which has nothing to do with how large k is made, and the consequent approximation $\sin(\delta/2) = \delta/2$, that make Madhava's interpolation a useful calculational tool.

All the same, there must have been a certain temptation to let k tend to infinity; Shankara almost does so. The argument of the last term, $\theta+\delta/2^k$, becomes just θ in the limit and we have an exact (at least formally) infinite series representation whose terms are $\sin(\delta/2), \sin(\delta/2)\sin(\delta/4), \cdots$, with coefficients linear in $\sin\theta$ and $\cos\theta$ alternately. It is a rather remarkable series, but it is definitely not a power series in δ and not very handy for any analytical purpose. We can see why Jyeshthadeva is so dismissively short in his treatment of interpolation; its value is purely computational, in which role it is very effective for angles not too close to zero.

If we set $\sin(\delta/2^i) = \delta/2^i$ for all i, we do get a power series in δ (which we may call the interpolation series), with coefficients linear alternately in $\sin\theta$

and $\cos\theta$, of which the first few terms are

$$\sin(\theta+\delta)_{inter} = \sin\theta + \delta\cos\theta - \frac{\delta^2}{2}\sin\theta - \frac{\delta^3}{8}\cos\theta + \cdots$$

and

$$\cos(\theta+\delta)_{inter} = \cos\theta - \delta\sin\theta - \frac{\delta^2}{2}\cos\theta + \frac{\delta^3}{8}\sin\theta + \cdots .$$

But it is a pointless thing to do. Numerically, for a given division of the quadrant, δ is bounded by the step size and one only has to carry out the iteration to the point where the omitted correction term is comparable to the error in the approximation $\sin(\delta/2) = \delta/2$; after that it is wasted labour. Analytically, the power series will not equal (will not converge to) $\sin(\theta+\delta)$. They do not coincide at $\theta = 0$ with Madhava's series; at an even more elementary level, without getting into the modern theory of power series, these strange series cannot represent $\sin\delta$ and $\cos\delta$ since they are not each other's derivatives, as Nilakantha and Jyeshthadeva (who both used the interpolation in their astronomy) knew very well they should be. The point is that the method of iterative refining is often accompanied by other approximations, specific to the application in mind, whose reliability is to be assessed independently of the refining process itself.

If I have given more space to the interpolation formulae than *Yuktibhāṣā* itself does, it is partly because the interpolation series and their superficial similarities with the Taylor series for the same functions

$$\sin(\theta+\delta) = \sin\theta + \delta\cos\theta - \frac{\delta^2}{2!}\sin\theta - \frac{\delta^3}{3!}\cos\theta + \cdots$$

and

$$\cos(\theta+\delta) = \cos\theta - \delta\sin\theta - \frac{\delta^2}{2!}\cos\theta + \frac{\delta^3}{3!}\sin\theta + \cdots$$

have recently been taken seriously as being representative of the infinitesimal mathematics of the Nila school[3]: the signs of successive terms match as do the alternating $\sin\theta$ and $\cos\theta$ in the coefficients; even the numerical coefficients are the same in the first three terms. The Taylor series for sin and cos are among the earliest of the triumphs of European calculus. Consequently, the interpolation series, raised in status from that of an unjustified extension of an approximation scheme to that of a "Taylor-like" series, became a touchstone in an assessment of whether Madhava's breakthroughs really qualified as calculus. This is an issue we shall revisit later, after a detailed study of Madhava's (correct) sine and cosine series. For the present it is sufficient to stress again that, as power series for $\sin(\theta+\delta)$, the interpolation series are incorrect: $\sin(\theta+\delta)_{inter} \neq$

[3]See Kim Plofker, "Relations between Approximations to the Sine in Kerala Mathematics" in [CHIM] and several of the references therein. The negative impact the misunderstanding had on a proper conceptual evaluation of the infinitesimal methods of Madhava ("Was calculus invented in India?") can also be traced from these references.

$\sin(\theta + \delta)$; and they are entirely irrelevant to the calculus credentials of the Nila school.

10.4 Saṃskāram: Generating Infinite Series

The word *saṃskāram* literally means 'refining' and comes from Sanskrit in both senses – it is taken from that language and the name *saṃskṛta* of the language itself is the past participle of its verb form. It, or expressions with the same functional meaning (e.g., *śuddhi, śodhya*, etc., used in the same sense already in Bakhshali), occur frequently in *Yuktibhāṣā* and in the other (Sanskrit) Nila texts to connote a process of making a gross (*sthūlam*) result more accurate (*sūkṣmam*) but also stands for its repeated use in the search for finer and finer estimates. There are variations in the way it is employed but there is an underlying unity of approach as well. The earliest example of one variant is the Bakhshali square root algorithm; one makes an informed first guess of $\sqrt{n}$ as the square root of the nearest square m^2 and – in a slight reformulation of the steps involved – sets up an equation for the difference r between the unknown $\sqrt{n}$ and its chosen approximate value m. This is a quadratic equation, $n = (m+r)^2$, whose solution for r involves a square root. The method outlined in Chapter 6.5 can then be restated as follows: evaluate the square root by the same ansatz, i.e., as the square root of the nearest square plus a correction, use it to find an approximate r by linearisation and compute the new error $r' = \sqrt{n} - (m + r)$ and repeat. After every step, the resulting equation is an identity because it only defines the last correction; an approximate and computable result emerges when the correction after a finite number of steps is replaced by its linear approximation. Alternatively, continue the process *ad infinitum*. If the initial guess is not too wild, successive corrections will become smaller and smaller and we end up with an infinite series, each term of which is easily computed (no square roots), that will converge, hopefully, to the sought-for exact answer. It actually does converge for the choice of m^2 as the square nearest to n even when n is as small as 2 (the *Śulbasūtra* square root). From the computational angle, there is thus a trade-off between an algorithm (the square root) which we do not know how to implement and an infinite sequence of easily executed steps. The exact series of course has an analytic significance beyond the algorithmic. The square root seems to be the only function explicitly to have been evaluated in this way in the pre-Nila days – though there are a fair number of instances of recursive reasoning – and its exact infinite series representation seems never to have been considered.

From the *saṃskāram* point of view, the square root problem is that of determining the value of the square root function at a point $m^2 + r$ given its value m at m^2. The interpolation of the last section is different only in detail, not in its philosophy. Ignoring for the moment the complication introduced by the cross-refining, the problem is of the same general kind: determine the value of a function f at a point $x + y$ knowing its value at x (from the sine table,

just as the square root function is known at all perfect squares) and knowing also that the difference $f(x+y)-f(x)$ is proportional to the same function at another point $x+y/2$, $f(x+y)=f(x)+\alpha(y/2)f(x+y/2)$ for a function α. Iteration is straightforward: we have

$$f(x+y)=f(x)+\alpha\left(\frac{y}{2}\right)\left(f(x)+\cdots\alpha\left(\frac{y}{2^{k-1}}\right)\left(f(x)+\alpha\left(\frac{y}{2^k}\right)f\left(x+\frac{y}{2^k}\right)\right)\cdots\right).$$

Once again we have the choice of terminating the iteration by setting $f(x+y/2^k)=f(x)$ for some k depending on the acceptable degree of grossness, or of continuing indefinitely to arrive at an infinite series of the form

$$f(x+y)=f(x)\left(1+\alpha\left(\frac{y}{2}\right)+\alpha\left(\frac{y}{2}\right)\alpha\left(\frac{y}{4}\right)+\cdots\right);$$

whether the series converges or not depends on the function α. Slightly more generally, the same conclusions follow if $f(x+y)=f(x)+\alpha(y)f(x+y_1)$ and the sequence $y_1,y_2\cdots$ converges to 0. Finally, suppose we have two functions f and g such that $f(x+y)=f(x)+\alpha(y)g(x+y_1)$ and $g(x+y)=g(x)+\beta(y)f(x+y_1)$. Then we can expand both f and g in infinite series by cross-refining and that is the situation for the interpolation series for sin and cos (with $\alpha=-\beta$ and other appropriate choices). However, as already emphasised in the previous section, any approximation made in the function α (such as the replacement $\sin(\delta/2)\to\delta/2$ in the interpolation formulae) will have an effect on the convergence properties and the resulting series will not in general equal $f(x+y)$.

The main reason I have thought it necessary to go over the formal structure of what we may call the general *saṃskāram* series is that it will be seen to underpin virtually every instance of the occurrence of infinite series in the Nila body of work; indeed it acts as a natural and unifying principle for the generation of infinite series expansions of functions, through the recursive application of the four fundamental operations of arithmetic/algebra. It is distinct from the production of Taylor series expansions of 'good' functions of a real variable and is logically independent of its principal ingredient, namely, infinitesimal calculus – we have seen for instance that it works just as well for the square root function on positive integers without embedding the integers in the real line. That is of course as true of the European marriage of the basic ideas of calculus with infinite series as of the Indian, though it is sometimes forgotten, perhaps because in Europe the two were born together. Unending (*ananta*) sequences and processes were part of the Indian mathematical world from the time it was realised that there was no end to numbers, certainly by the time of the lists of powers of 10 (see Chapter 5.3). In Europe the realisation came late, with the spread of decimal arithmetic. Here is a long and revelatory passage from Newton's writings on infinite series (1670-71), specifically power series (in Whiteside's translation from Latin; the italics are mine)[4]:

[4] *De Methodis Serierum et Fluxionum*, in [Ne-W], vol. III, p.32

> Since the operations of computing in numbers and with variables are closely similar – indeed there appears to be no difference between them except in the characters by which quantities are denoted, definitely in the one case, indefinitely so in the latter – I am amazed that it has occurred to no one (if you except N. Mercator with his quadrature of the hyperbola) to fit the *doctrine recently established for decimal numbers* in similar fashion to variables, especially since the way is then open to more striking consequences. For since this doctrine in species has the same relationship to Algebra that the doctrine in decimal numbers has to Arithmetic, its operations of Addition, Subtraction, Multiplication, Division and Root-extraction may easily be learned from the latter's provided the readers to be skilled in each, both Arithmetic and Algebra, and appreciate the correspondence between decimal numbers and algebraic terms *continued to infinity*; namely that to each single place in a decimal sequence decreasing continually to the right there corresponds a unique term in a variable array ordered according to the sequence of the dimension (the 'place' in a decimal number, the degree of the variable in one term of a polynomial; the word 'dimension' comes from the identification of the kth power with the volume of the k-dimensional cube) of numbers or denominators.

One can of course parse the passage *in extenso* and derive much historical insight in the process, but two facts stand out, quite apart from the impact the "new doctrine" of decimal enumeration and arithmetic had on European mathematical thinking. It marks a coming to terms with power series as a respectable mathematical object and it is even more remarkable for the explicit parallel it draws between algebra and arithmetic, impossible to envisage without place value numbers; Brahmagupta might have written it if he was given to such an expansive style of writing.

Not only the acceptance of infinite series as a precisely defined mathematical object, but also the process itself of generating successive terms by feeding the output of one step as the input in the next step – recursive construction in a broad sense – derives ultimately from that most fundamental of recursive constructs, based natural numbers in their infinitude. We will return to the role of recursive reasoning and recursive constructions as one of the defining characteristics of Indian mathematics at the end of this book.[5] For the present, I close this section with a particularly simple example that Nilakantha uses (in the *Āryabhaṭīyabhāṣya*) to explain some of the theoretical issues that come up in applications of the method of recursive refining. As it turns out, it is also an indispensable part of the calculus of the π series where its value is not in any sense algorithmic but because it makes possible, ultimately, the integration of

[5]In the meantime readers interested in exploring the theme beyond the algorithmic use of *saṃskāram* and its formal basis described here may find my article: P. P. Divakaran, "Notes on *Yuktibhāṣā*: Recursive Methods in Indian Mathematics" in [SHIM] useful.

a certain function – a precedent which Newton too, independently, followed. It is formulated in superficially different ways in *Āryabhaṭīyabhāṣya* and *Yuktibhāṣā*. The account below covers and is faithful to both; the only liberty I have taken is to make the notation and the sequencing of steps conform to the approach adopted in this section to *saṃskāram* as a method of solving functional equations.

The function that is to be 'refined', i.e., rewritten as an infinite series, is one that cannot be considered 'unknown' by any stretch of the imagination, namely $f(x) = 1/x$. To determine $f(x+y)$, write it in the by-now conventional manner as

$$\frac{1}{x+y} = \frac{1}{x} - \left(\frac{1}{x} - \frac{1}{x+y}\right) = \frac{1}{x} - \frac{y}{x} \cdot \frac{1}{x+y},$$

displaying the supposedly unknown quantity as a factor in the correction term on the right, suitable for further refining:

$$\frac{1}{x+y} = \frac{1}{x} - \frac{y}{x}\left(\frac{1}{x} - \frac{y}{x}\left(\frac{1}{x} - \cdots\right)\right).$$

After k steps, we have the identity

$$\frac{1}{x+y} = \frac{1}{x} - \frac{y}{x^2} + \frac{y^2}{x^3} - \cdots + (-1)^k \frac{y^k}{x^k} \cdot \frac{1}{x+y}$$

valid for all k and all values of x and y. Putting $y = 0$ in the last factor in the last term on the right is the equivalent here of making the unknown function approximately computable, though everything is known from the beginning: no one would compute $1/(x+y)$ by *saṃskāram*. As Nilakantha makes very clear, the interest of the example is fundamentally conceptual and theoretical.

Accordingly, take the limit $k \to \infty$: we end up with the formal geometric series representation

$$\frac{1}{x+y} = \frac{1}{x}\left(1 - \frac{y}{x} + \frac{y^2}{x^2} - \cdots\right).$$

Whether the series on the right equals $1/(x+y)$ (or any function of x and y for that matter) depends on whether it is a convergent series, i.e., on the condition $|y/x| < 1$ being satisfied. Nilakantha actually considers the series $1/(x-y) = (1/x)\sum_{k=0}^{\infty}(y/x)^k$, characterises it as an infinite geometric series and says explicitly that the common ratio x/y should be greater than 1, illustrating the reasoning for the numerical choice $x = 4, y = 1$ – i.e., the series $1/3 = \sum_1^{\infty}(1/4^k)$ – by writing out the first few steps in the *saṃskāram*. Even more pointedly, he asks at the beginning of his explanation how we can be sure that the infinite series thus generated will attain but not exceed $1/(x-y)$, conveying thereby that the series should converge for it to make sense and, when it does, that the value of $1/(x-y)$ is attained as the limit of an infinite sequence of numbers, each of which is the output of a finite algorithm. The question of convergence must have been a preoccupation; to anticipate a little,

Yuktibhāṣā has a careful statement to the effect that the general arctangent series, θ in powers of $\tan\theta$ for $0 \le \theta \le \pi/4$, cannot hold in the second octant $\pi/4 < \theta \le \pi/2$ because the variable $\sin\theta/\cos\theta$ is then greater than 1 and that the correct variable to expand in then is $\cos\theta/\sin\theta$.

Two general remarks. As we know, the geometric series can be looked at as the specialisations of the binomial series for exponent -1 or, more generally, of the Taylor series. The general binomial series for an exponent not a positive integer was not an object of study in the Nila school – or earlier – perhaps because it was not needed. The absence of the Taylor series is more puzzling since, for trigonometric functions, it is only a short step from Madhava's sine and cosine series (which are after all their Taylor series at $\theta = 0$) and the addition theorems (see below, Chapter 12.4). If nothing else, that would have helped put the 'Taylor-like' interpolation series in its own proper setting.

We may also wonder, in retrospect, why the concern about limits of infinite sequences of numbers was not extended to other similar contexts, for example the value of π from the perimeter of Aryabhata's inscribed regular hexagon (or the Nila school's circumscribing square) doubled indefinitely, a would-be Nila version of Archimedes' method of exhaustion so to say. The greater complexity of the problem would have made it a challenge but harder challenges had been met and overcome. It is possible that the conjectured irrationality of π acted as an inhibitor: if "the true value [of π] cannot be given" (Nilakantha), how can it be expressed as the limit of a sequence of numbers, involving multiple square roots, which themselves "cannot be determined" according to the same Nilakantha? Perhaps *Yuktibhāṣā*'s stated motivation for the π series, "a method of producing the circumference without taking square roots", reflects this dilemma. It seems reasonable to conclude that Nilakantha's clarification of what he meant by "attains but does not exceed", in other words the notion of a limit, was provoked by the need to convince himself that the π series does converge to π even if its "true value cannot be given".

11

The π Series

11.1 Calculus and the Gregory-Leibniz Series

In the context of its European history, the question of what exactly the branch of mathematics called calculus consists of and what its foundations are is an old and much debated one, raised by mathematicians and philosophers alike already while Newton and Leibniz were alive. The issue was not settled, to the satisfaction of at least the mathematicians, until the 19th century. Little that is worthwhile can be added anew to that discussion. But, given that there are interesting differences in the Indian and European approaches to the "metaphysics of calculus" – a philosophically evocative phrase first used by D'Alembert in the 18th century – even as they are in agreement on the invariable core idea, at least a brief reexamination of some of these issues, in their most basic manifestations, will become unavoidable. The best way to do that seems to be to accompany or interpose the descriptions of the calculus-related material in the Nila work – elementary from our present viewpoint – with the corresponding material as treated in the classrooms of today.

The first such problem is the computation of π as the constant of proportionality relating the circumference of a circle to its diameter or, more generally, the length of an arc of the unit circle as a function of its half-chord, resulting respectively in the basic π series ('basic' to distinguish it from other, more rapidly converging, series for π also due to Madhava, which are derived from the basic series) and the arctangent series. It is the first problem in more than the chronological sense. In the Nila work, it is the problem for which calculus was invented. In Europe also, it is the simplest example of one of the driving forces behind the invention of calculus, namely the rectification of a curve: the determination – as well as a satisfactory definition – of the length of a (plane) curve, measured along the curve. More than that, the solution of the rectification problem expresses most directly the inverse relationship between the operations of differentiation and integration (the fundamental theorem of calculus) and in fact led to the first formulations of the latter.

P. P. Divakaran, *The Mathematics of India*, Sources and Studies in the History of Mathematics and Physical Sciences, https://doi.org/10.1007/978-981-13-1774-3_11

The term 'Gregory-Leibniz series' in the heading of this section covers both the general series

$$\theta = t - \frac{t^3}{3} + \frac{t^5}{5} - \frac{t^7}{7} + \cdots ,$$

where $t = \tan\theta$, and its special case for $\theta = \pi/4$,

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots ,$$

specifically in the context of their European history. I start by giving a typical modern derivation of the result, partly in order to pinpoint the ingredients – in particular the calculus-related ideas – that go into it and partly as a benchmark against which to compare the derivation in *Yuktibhāṣā*. To determine θ as a function of t, restrict t to the interval $0 \leq t \leq 1$ and differentiate $\tan\theta = \sin\theta/\cos\theta$ with respect to θ. The answer, either directly from geometry, or from the formula for derivatives of rational functions together with the derivatives of sin and cos (geometry again), is $dt/d\theta = 1+t^2$. Invert it and expand $(1+t^2)^{-1}$ in a power series (as a binomial series or from the formula for the sum of a geometric series):

$$\frac{d\theta}{dt} = \frac{1}{1+t^2} = 1 - t^2 + t^4 - t^6 + \cdots$$

and integrate term by term using the formula

$$\int_0^t u^k du = \frac{t^{k+1}}{k+1}$$

for $k = 0, 2, 4, \cdots$.

The steps in this proof are all standard and their very familiariy is reason enough to remind ourselves where they come from (the references to the Nila work are, for the moment, meant only to evoke what we have already seen or in anticipation of what is to come):

1. The differentiation is essentially geometry in the small: find the tangent at a point θ to the graph of the function $\tan(\theta)$. It can be approached in apparently different ways, Leibniz's original method and the *Yuktibhāṣā* method to be described below being examples, but they all boil down to the geometric properties of the circle as the lengths of relevant chords are made smaller and smaller.

2. Inverting the derivative (effectively, making t the independent variable) is a natural operation if it is understood that the aim of the exercise is to determine θ as a function of t and, eventually, to rectify an octant of the circle (as it was in the Nila work). Since $d\theta/dt$ comes out as an explicit function of t, it also, conveniently, reduces the subsequent integration to a quadrature. The question of the expressibility of the derivative as a

function of the independent variable was an important one historically, for instance in Newton's derivation of the sine series, as we shall see later on.

3. But the quadrature does not lead to a simple function of t. There is a good reason why it cannot and why one is led to the series expansion: if the integration were to result in a rational function of t (with rational coefficients) for instance, then the special case $t = 1$ would give a rational value to π. The same power series expansion of $(1 + t^2)^{-1}$ is resorted to already in the Nila treatment of the problem and we can wonder whether such thoughts might have been behind Nilakantha's conviction that π is irrational or, conversely, whether the irrationality might have pointed to the need for the infinite series expansion.

4. The series expansion works because we have prior knowledge of the integrals of positive integral powers. The knowledge comes from knowing the derivatives of powers and the (implicit) use of the fundamental theorem, as it did for the pioneers of European calculus. The Nila mathematicians did not work out the derivatives of powers – they could have done so if they wished by looking at deviations from the rule of three – and so had to compute the power integrals from first principles as limits of sums. Several originalities make their first appearance in this part of the work: auxiliary results like discrete integration by parts, new methods of proof such as mathematical induction, etc.

5. A modern account of the π series will also include a proof that it is convergent (but not absolutely convergent). In the early days of European calculus, the idea of infinite series as respectable mathematical objects was still new (see the quotation from Newton in Chapter 10.4 and footnote 9.4) and notions relating to convergence very fuzzy. The textbook proof of convergence proceeds by writing $(1+t^2)^{-1}$ as the sum of the first k terms and a remainder (identical with Nilakantha's *saṃskāram* identity for the geometric series, Chapter 10.4) and noting that the integral of the remainder vanishes in the limit $k \to \infty$. Despite Nilakantha's concern about the sum of the geometric series attaining but not exceeding its limiting value, no Nila text refers explicitly to the convergence properties of the π series itself. But, as noted in Chapter 10.4, *Yuktibhāṣā* does remark that the appropriate variable in which to expand the arc θ when it extends into the second octant, where $\sin\theta > \cos\theta$, is not $\tan\theta$ but its inverse.

6. Integration is an arithmetical or algebraic process, that of adding up the arc segments in the limit in which they tend individually to 0. In the European metaphysics of calculus, the limit is effectively taken before the sum by defining the differential (Leibniz) or, equivalently, the derivative or fluxion (Newton). (This of course was the problematic step). Correspondingly, the sum becomes the integral in the limit. The cornerstone of the entire procedure is therefore the fundamental theorem of calculus

which asserts, in its simplest version for a real function f of a real variable x, the equality

$$\int_a^b df = \int_a^b \frac{df}{dx} dx = f(b) - f(a).$$

In the Nila work, in contrast, the limit is taken after the sum is computed: what is added up are the finite arc segments (in the problem at hand; more generally, finite differences of appropriate geometrically defined functions). As noted at the end of Chapter 7.5 in connection with Aryabhata's use of finite differences and discrete integration in the preparation of the sine table, the fundamental theorem is a triviality before the limit is taken. But there is a price to pay: the exact finite differences cannot in general (for example in the present problem of the rectification of the circle) be summed up in a simple closed expression; they have to be suitably approximated and the resulting error shown to vanish in the limit.

The question of the timing of the limit is also relevant to the inversion of the derivative (point 2 above). While the reciprocal relationship is valid for the ratio of finite differences δt and $\delta\theta$, it needs to be proved that it survives the limit and continues to hold for the derivative. It is not an issue in the Nila approach where all arithmetical operations are completed ahead of the final limit operation.

As far as their European setting is concerned, the points made above have become part of the foundations of calculus as they evolved from the 17th century onwards in order to accommodate the many directions in which the discipline has been generalised. The secondary comments on how the corresponding ideas in the Nila approach compare with these standard notions may appear sketchy to readers not familiar with the Nila work; their relevance will become clearer as we work through the details.

11.2 The Geometry of Small Angles and their Tangents

As befits the first ever illustration of how the new infinitesimal ideas are put to use, *Yuktibhāṣā*'s account of the π series is very detailed and careful. As elsewhere in this book, I have adapted its narrative style to modern conventions, notation and terminology (and rearranged the order of presentation here and there) while keeping faith with the concepts and their detailed working out.

Draw the square in which the unit circle, with O as centre, is inscribed. If A is the point of contact of the circle with one of the sides and B the vertex of the square to the right of A (conventionally, OA is the direction east and B is in the direction southeast), then the arc between OA and OB is one-eighth of the circumference, subtending an angle $\pi/4$ at O. The line AB is tangent to the circle at A (perpendicular to the radius OA) and of unit length. If Y is a

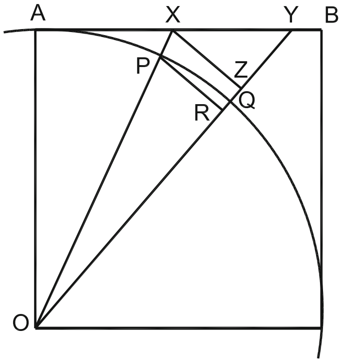

Figure 11.1: The geometry of the tangent.

point on AB with angle$(OA, OY) = \theta$, then AY is $\tan\theta = t$. Geometrically, what is required is the arc AQ, where Q is the intersection of OY with the circle, as a function of the line AY: in the Nila approach, t is the independent variable from the beginning.

Divide now the line AB into a large number n of equal parts by points A_i, with $A_0 = A$ and $A_n = B$. Identify Y with one of the points $A_i (i \neq n)$ and denote A_{i-1} by X as in Figure 11.1 (on a highly magnified scale, for ease of drawing). Join O to X and let OX intersect the circle at P. Then t is i/n, the small line segment $XY = 1/n$ is the increment δt, independent of i, in the tangent and the small arc segment PQ is the corresponding change $\delta\theta_i = \theta_i - \theta_{i-1} = \text{angle}(OX, OY)$ in the angle, the subscript i being the finite-difference indicator of the t dependence of $\delta\theta$. The construction is completed by drawing the perpendiculars PR and XZ to OY. PR is obviously $\sin\delta\theta_i$; XZ has no fundamental geometric significance but plays a useful intermediate role in the reasoning.

The method of calculus requires first the change $\delta\theta_i$ in the arc (arc PQ) to be expressed in terms of $\delta t (= XY)$. *Yuktibhāṣā* does not do so directly; it finds instead a relation between XY and $PR = \sin\delta\theta_i$, in the knowledge that eventually $\sin\delta\theta_i$ will be replaced by $\delta\theta_i$ as n is made larger and larger. The geometric steps needed are, as always, reduced to an astute choice of similar triangles (which is where XZ comes in). The right triangles XYZ and OYA are similar because the non-right angle at Y is common to both as we would argue today. (*Yuktibhāṣā* applies the criterion that the sides of two right triangles obey the rule of three if the hypotenuse of either triangle is parallel to a short side of the other; this is a criterion not met with before and so is explained in detail with an illustration from architecture). Consequently, we have

$$\frac{XY}{OY} = \frac{XZ}{OA} = XZ$$

since $OA = 1$. Next, from the pair OXZ and OPR of similar right triangles (this is simple *Śulbasūtra* scaling), we also have

$$\frac{XZ}{PR} = \frac{OX}{OP} = OX$$

since $OP = 1$. Eliminating the uninteresting XZ, we get the basic relation

$$PR = \frac{XY}{OX \cdot OY}.$$

As is evident from the derivation, this relation holds for any two points X and Y on AB, free from assumptions about the eventual smallness of XY. Going back to the identifications $X = A_{i-1}$, $Y = A_i$, $AX = (i-1)/n$, $AY = i/n = t$, $\delta t = 1/n$, arc $PQ = \delta\theta_i$ and $PR = \sin\delta\theta_i$, we can rewrite it, again exactly, as

$$\sin\delta\theta_i = \frac{\delta t}{d_{i-1}d_i} = \frac{1}{nd_{i-1}d_i},$$

where d_i is the length of the diagonal (*kaṛṇṇam*) joining O to $A_i = Y$.

The diagonals can now be expressed in terms of the tangents: $d_i^2 = OY^2 = OA^2 + AY^2 = 1 + i^2/n^2$. But before doing so, *Yuktibhāṣā* makes a first use of the limit of large n through the approximation

$$\frac{1}{d_{i-1}d_i} = \frac{1}{2d_{i-1}^2} + \frac{1}{2d_i^2}.$$

No justification is offered for the asymptotic ($n \to \infty$) exactness of the approximation beyond the statement "the two [adjacent] diagonals are almost equal in magnitude" but we can easily supply one. The difference of the two sides of the approximate equation is easily computed to be $(d_i - d_{i-1})^2/(2d_{i-1}^2 d_i^2)$; since $d_i - d_{i-1}$ is of order $1/n$, the error is of the second order of smallness and the contribution of n such terms will be vanishingly small when added together in the integration to follow. Indeed, while doing the integration, *Yuktibhāṣā* makes a second asymptotically exact approximation, replacing d_{i-1}^2 by d_i^2 in the denominator; the resulting simple expression for the half-chord corresponding to the ith tangent segment of length δt is

$$\sin\delta\theta_i = \frac{\delta t}{d_i^2} = \frac{1}{n(1 + i^2/n^2)}$$

and its justification is given in the text: the sums over the terms indexed by $i-1$ and i coincide except for the first term $1/d_0^2$ $(= 1/OA^2 = 1)$ in the former and the last term $1/d_n^2$ $(= 1/OB^2 = 1/2)$ in the latter; they are finite and can be neglected when divided by $n \to \infty$.

The crucial step now is to equate the half-chord $\sin\delta\theta_i$ to the arc-segment $\delta\theta_i$, again exact in the asymptotic limit: "If the segments of the [half] side of

the square (δt) are extremely small, these half-chords are the same, very closely, as the [corresponding] arc-segments" (6.3.1).[1] We have thus the expression

$$\delta\theta_i = \frac{1}{n(1 + i^2/n^2)}$$

for the change in the angle when its tangent changes by $1/n$, exact in the limit $n \to \infty$. *Yuktibhāṣā* stops here for the time being, since a final limit is still to be taken after the sum over i is performed (integration). But there is nothing to stop us from dispensing temporarily with the index i in favour of the continuous variable t:

$$\delta\theta(t) = \frac{\delta t}{1 + t^2}$$

except that we have to worry, as Europe did, about the meaning to be given to the vanishingly small, 'infinitesimal' (*aṇuprāyam*, of the nature of an atom, in *Yuktibhāṣā*'s words), quantities δt and $\delta\theta$. In a modern reading they are, obviously, to be interpreted as differentials.

It is no surprise that *Yuktibhāṣā*'s solution of the differential part of the problem of rectifying (an arc of) the circle is the standard one we are familiar with, expressed in exactly the same formula that results from the standard rules of differential calculus. Perhaps less expectedly (but as advertised), the principle behind the result and the method employed to arrive at it are also seen to be essentially no different. First linearise the arc locally, not by the line connecting its ends (the chord) but, following Aryabhata's lead, by half the chord of twice the arc (the sine). Then take the limit in which the arc tends to zero; in the limit it does not matter whether the linearisation is done through the chord or the half-chord since they both tend geometrically to the tangent (of infinitesimal length) to the infinitesimal arc. And, within the Indian context too, this step marked a conceptual liberation: the ghost of the fallacy of Bhaskara I, that there is a non-zero angle for which the chord equals the arc (Chapter 7.5), was finally laid to rest; it had held its own for almost a millennium, coming in the way of an understanding of infinitesimal quantities – even by someone as meticulous in his thinking as Bhaskara II, as we saw in connection with division by zero or infinity (Chapter 5.4) or in the computation of the surface area and volume of the sphere (Chapter 7.5). It is particularly striking, and something of an irony, that the key to a proper handling of the infinitesimal was an idea that both Bhaskara and Brahmagupta struggled with, the reciprocal relationship between zero and infinity, reexpressed in terms of division by unboundedly large numbers.

---

[1]Unattributed quotations in this and the following chapters are from *Yuktibhāṣā*, in my translation. The numbering is that in K. V. Sarma's book [YB-S]: N.n.m refers to chapter N, section n and subsection m (if there is one).

11.3 Integration: The Power Series

Yuktibhāṣā describes every step in the summation of the linearised arc segments

$$I_n = \sum_{i=1}^{n} \sin \delta\theta_i = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{1 + i^2/n^2}$$

and the passage to the limit

$$I = \lim_{n \to \infty} I_n$$

in elaborate detail. Considering how many of the ideas involved were entirely new to the mathematical culture of the times, it had good reasons for doing so. It is not necessary or practical to go over all of it here in equal detail; we have met some of the steps earlier and others involve operations that have become standard since then and are very familiar to a modern student. The account in the present section is therefore somewhat streamlined[2]: in particular the occasionally tedious case-by-case analyses – the price paid for the absence of an efficient notation – of some of the novel points that come up are dealt with below directly in the general case, without deviating, it is hoped, from the logical and mathematical line of thought of the text.

Preliminary to the summation is the expansion of $(1 + i^2/n^2)^{-1}$:

$$\frac{1}{1 + i^2/n^2} = 1 - \frac{i^2}{n^2} + \frac{i^4}{n^4} - \cdots .$$

The demonstration is the same as the *saṃskāram* of Nilakantha (see Chapter 10.4) though the terminology is slightly different: successive terms are called *śodhyaphalam*, literally "the result of purification", reminiscent of the terminology of the Bakhshali manuscript in the recursive refining of the square root. Along the way, the question of the negligibility of the error after a large but finite number of terms, convergence in modern parlance, is addressed and illustrated with a numerical example. Disappointingly, no motivation for resorting to the expansion is provided, nothing about the difficulty in integrating $(1 + t^2)^{-1}$ or the irrationality of π.

The resulting infinite series for I_n is thus

$$I_n = I_{n,0} - I_{n,2} + I_{n,4} - \cdots ,$$

with

$$I_{n,k} = \frac{1}{n^{k+1}} \sum_{i=1}^{n} i^k$$

for $k = 0, 2, 4, \cdots$. It is assumed in this step that the (infinite) sum over k and the (as of now finite) sum over i can be interchanged; this is perhaps one of

[2] A reading of the relevant sections, 6.3.3 to 6.4.4 primarily, of Sarma's translation ([YB-S]) is, nevertheless, very rewarding.

the reasons for insisting on the negligibility of the error in the k-summation after a large enough number of terms. Eventually, the limit $n \to \infty$ will have to be taken and it is another of the gains from postponing that limit until all computations are finished that we do not have to worry about the legitimacy of interchanging infinite sums and integrals.

The expansion reduces the problem to the evaluation of the quantities

$$J_k = \lim_{n\to\infty} I_{n,k} = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \frac{i^k}{n^k},$$

for $k = 0, 2, 4, \cdots$ and to adding them up with the apprpriate signs:

$$I = J_0 - J_2 + J_4 - \cdots .$$

In terms of the original continuous variable t with the identifications $t = i/n$ and $\delta t = 1/n$, J_k as the limit of the sum over i is precisely the modern definition of the integral of t^k:

$$J_k = \int_0^1 t^k dk.$$

I have compressed several of *Yuktibhāṣā*'s careful and precise explanations, both conceptual and methodological, but this in all essentials is how it reduces the problem finally to an infinite series of definite integrals of even powers. It is also not significantly different from the textbook reduction to integrals of powers as summarised in section 1 of this chapter, allowing, of course, for the fact that the notion of an integral was assumed there to be already acquired. The other, consequent, difference from the standard treatment is, as already noted, in the actual determination of the general power integral from first principles, as the asymptotic limit of a finite sum over i.

A small note of caution. *Yuktibhāṣā* uses the term *saṃkalitam*, with various qualifiers, for the sum (over i) whose limit is the integral (it is not used for the infinite sum over k) both before and after the limit, i.e., both for the discrete integral and the 'true' integral: thus *ghana-saṃkalitam* is the sum of cubes of natural numbers as well as the integral of the third power, etc., and *saṃkalita-saṃkalitam* is both a sum over sums and a multiple integral. Where appropriate, I will feel free to use the word 'integral' for *saṃkalitam* without further explanation.

The first surprise in the working out of the integrals of powers is that they are sought for all (non-negative) integral k, not just for the even k that occur in the expansion: " Even though it is not useful here, [I am] describing also the integrals of equals multiplied three, five, etc. times among themselves, as they occur in the midst of those which are useful" (6.4, opening paragraph). The reason for the broadening of the problem becomes evident in the following sections: J_k (or $I_{k,n}$ for large n) will be related to J_{k-1}; in other words, mathematical induction will play an indispensable part in the integration of the general power. In what is probably its most carefully written part, *Yuktibhāṣā*

describes the procedure in several steps. First J_1 is evaluated from J_0 and J_2 from J_1 with equal attention to detail; then J_3 from J_2 and J_4 from J_3 progressively more briefly. The general rule is then given: "To produce integrals of higher and higher powers, multiply the given integral by the radius and remove from it itself divided by the number which is one greater". (6.4.4). Altogether, it is difficult to escape the feeling that Jyeshthadeva is trying to convey to his disciples an unfamiliar and particulary subtle line of thought – recall his quasi-axiomatic treatment of based natural numbers through the property of succession (Chapter 4.2).

For the unit circle, the general inductive prescription above amounts to

$$J_k = J_{k-1} - \frac{J_{k-1}}{k+1} = \frac{k}{k+1} J_{k-1}$$

which implies by iteration (with $J_0 = 1$ as input)

$$J_k = \frac{1}{k+1}.$$

The *Yuktibhāṣā* proof of this fundamental result – the first ever example of the rigorous working out of a nontrivial integral – is both subtle and unexpectedly original to a modern student. Rather than follow it literally and describe first the special cases of small values of k individually I will, in the next section, do the case of general k, following the same arguments as used in *Yuktibhāṣā* but exploiting the flexibility and efficiency of present-day notation. Apart from saving space, it will also bring to light the modern analogues of some of what may at first sight appear to be just ingenious 'tricks'.

But, before that, two general remarks. The first is a historical-epistemic point we have encountered earlier (Chapter 7.1). The exact expressions for $I_{n,k}$ for $k = 1, 2, 3$ and for any n were known to Aryabhata (*Gaṇita* 19, 21 and 22) and were almost surely derived by geometric, 'building block', methods; at least that was how it was done in the Nila school if we go by Nilakantha in the *Āryabhaṭīyabhāṣya* (see Chapter 7.1). But Chapter 6 of *Yuktibhāṣā* does not refer to these exact low k results at all (except, implicitly, for $I_{n,0} = n$; $I_{n,1}$ is derived geometrically in Chapter 7) possibly because, as suggested by Sarasvati Amma in a related context, geometric imagination could not break through the barrier of the dimension of physical space. Inductive proofs seem to have been thought of as a substitute – from solid, down-to-earth, geometry to logical abstraction as it were. The idea would have found ready acceptance if not, necessarily, immediate comprehension; mathematical induction is after all the elevation to the status of a proof-device of another of the hallmarks of the Indian mathematical mind, the attachment to recursive reasoning.

The other remark is that induction is applied not to the exact sum of powers of natural numbers, but only to its dominant term in the large n approximation. Sums of general powers higher than 3 would not have been easy to work out (they involve the Bernoulli numbers). Working with the asymptotically dominant term not only met the needs of the problem at hand. The judicious neglect of subdominant terms also simplified the proof enormously.

11.4 Integrating Powers; Asymptotic Induction

What is required now is the asymptotic behaviour in n of the sums of powers of positive integers:

$$n^{k+1} I_{n,k} = S_n^k = \sum_{i=1}^{n} i^k,$$

where I have reverted to the notation of Chapter 7.1) for all k (keep in mind that the superscript k in S_n^k is not an exponent). Gathering together the common threads running through *Yuktibhāṣā*'s individual treatment of low values of k as well as its explanation of the general case, here is how it is done.

Replace one factor of i in i^k by n, thereby changing S_n^k to $n\sum_{i=1}^{n} i^{k-1} = nS_n^{k-1}$. It is natural to think of this substitution as the first guess in a process of *saṃskāram*: $S_n^k = nS_n^{k-1}+$ a (negative) correction to be determined; that is how *Yuktibhāṣā* proceeds (though without explicitly using the word *saṃskāram*). The error introduced by the substitution is

$$nS_n^{k-1} - S_n^k = \sum_{i=1}^{n-1} (n-i) i^{k-1},$$

the coefficient of the $i = n$ term being 0. The ingenious step now is to reorder the sum on the right as

$$\sum_{i=1}^{n-1} (n-i) i^{k-1} = \sum_{i=1}^{n-1} \sum_{j=1}^{i} j^{k-1} = \sum_{i=1}^{n-1} S_i^{k-1}.$$

The proof is a matter of enumeration of terms: the right side is, explicitly, $1^{k-1} + (1^{k-1} + 2^{k-1}) + \cdots + (1^{k-1} + 2^{k-1} + \cdots + (n-1)^{k-1})$; collecting coefficients of i^{k-1}, the expression becomes $(n-1)1^{k-1} + (n-2)2^{k-1} + \cdots + 1\cdot(n-1)^{k-1}$ which is the left side. The rearrangement has thus resulted in a recursion relation in k,

$$S_n^k = nS_n^{k-1} - \sum_{i=1}^{n-1} S_i^{k-1},$$

for the power sum.

In principle, one can iterate the step, reducing k by unity successively until $k = 0$ is reached. But the n-dependece coming from the second term on the right will be quite involved. *Yuktibhāṣā* chooses to circumvent such complications by taking advantage of the knowledge that n is eventually going to be made to tend to infinity and keeping only the dominant terms in the relevant n-dependent quantities. Start the induction with the trivial observation $S_n^0 = n$ for all n. We have then the recursion relation for $k = 1$:

$$S_n^1 = nS_n^0 - \sum_{i=1}^{n-1} S_i^0$$

in which the first term on the right is n^2 and the second term is $\sum_{i=1}^{n-1} i = S_{n-1}^1$:

$$S_n^1 = n^2 - S_{n-1}^1.$$

Now, for large n, ignore the difference ($= n$) between S_n^1 and S_{n-1}^1 (justified since S_n^1 increases quadratically with n). From the resulting approximate equation we get

$$S_n^1 \sim \frac{n^2}{2}$$

where, and in the following few equations, $\sim$ denotes the dominant term in the limit $n \to \infty$, including the coefficient.

Rather than repeat the steps above for $k = 2, 3$, and so on, let us incorporate them in a general inductive proof in the modern manner. Accordingly, suppose $S_n^{k-1} \sim n^k/k$ for a given $k > 0$. The first term on the right in the recursion relation is dominated by n^{k+1}/k. The crucial point in evaluating the second term is that it is legitimate to extend the induction ansatz to all i: $S_i^{k-1} \sim i^k/k$, when $n \to \infty$ and i is summed over. The reason is that the sum is dominated by terms corresponding to large n and the error introduced in the lower terms will sum to a finite quantity; stated otherwise, $\sum_{i=1}^{n-1} S_i^{k-1}$ is a polynomial one degree higher in n than S_{n-1}^{k-1}. The asymptotically dominant part of the second term is therefore

$$\sum_{i=1}^{n-1} S_i^{k-1} \sim \sum_{i=1}^{n-1} \frac{i^k}{k} = \frac{1}{k} S_{n-1}^k \sim \frac{1}{k} S_n^k,$$

where we have once again ignored the difference between S_n^k and S_{n-1}^k ($= n^k$). The recursion relation then gives, for the dominant terms,

$$\frac{k+1}{k} S_n^k \sim \frac{n^{k+1}}{k}$$

or

$$S_n^k \sim \frac{n^{k+1}}{k+1}$$

which is what is needed to be shown. Alternatively, we can leave the first term as it is and determine the asymptotic values of S_n^k recursively:

$$S_n^k \sim n \frac{k}{1+k} S_n^{k-1}$$

which is what *Yuktibhāṣā*, not surprisingly, prefers.

Finally, for the quantities $I_{n,k}$ themselves (whose limits are the integrals of powers) we have

$$I_{n,k} = \frac{1}{n^{k+1}} S_n^k \sim \frac{1}{k+1}$$

which is the same as

$$J_k = \int_0^1 t^k dt = \lim_{n \to \infty} I_{n,k} = \frac{1}{k+1}.$$

The rectification of the (octant of the) circle is complete:

$$\frac{\pi}{4} = I = J_0 - J_2 + J_4 - \cdots = 1 - \frac{1}{3} + \frac{1}{5} - \cdots .$$

It bears repetition that every step in my account of *Yuktibhāṣā*'s progression from the initial segmentation of the unit tangent to the final series is as it is in the book. What it may have missed in the transcription to today's terminology and notation is the strong overall sense it conveys that it is treading here on uncharted territory. There are places where we can get a glimpse of the struggle it must have been to come to terms with the novelty of some of these steps, for example the justification of the asymptotic ansatz for $\sum_i S_i^k$ when the sum is taken to infinity, or in the repeated description of the inductive steps from $k = 1$ to $k = 4$, as though one needed to be triply convinced of the acceptability of formal induction as a method of proof. There is also an instance of a false justification. After establishing the π series, with its alternating terms progressively diminishing in magnitude (a point it draws attention to), *Yuktibhāṣā* says: "Since the numerator is less than the denominator and the result (the quotient) is successively smaller, when it becomes extremely small we can omit the subsequent results and terminate the computation"(6.4.5). We know now that "we can omit the subsequent results" because it is also an alternating series, not just a series in which the terms are successively smaller. The historically interesting point here is not so much that subtelities regarding the notion of conditional convergence were not grasped – it would be astonishing if they were – as that the issue of convergence was given serious thought at all.

In contrast, no such indications of insecurity are visible – even though the account is occasionally laboured and repetitive – when it comes to the heart of the matter, the limiting operation itself (and the arguments for the need for it in arriving at the exact answer). *Yuktibhāṣā* uses 10^{17} (*parārddham*; it is the largest named power of 10 in its list) for the large number n into which the tangent is divided but adds that it is only a notional number which must eventually be made unboundedly large to produce the sum of the infinitesimally small segments. The procedure is simple and straightforward once the idea is grasped: the geometric quantity which is the independent variable (the tangent in the present problem) is divided into n equal parts and the desired dependent quantity (the arc) is computed for each segment dropping all subdominant contributions, the whole to be added up and the limit $n \to \infty$ taken at the end. (We shall see in due course that exactly the same logical sequence of steps is followed in the other calculus problems that *Yuktibhāṣā* deals with). Rather than having to define infinitesimal quantities *ab initio* as in Europe, this way of dealing with them through standard arithmetical/algebraic procedures followed by one limiting operation – division by infinity in short – leads to a metaphysics of calculus that is free at all stages from mathematical and logical ambiguity.

The postponement of the division by infinity until the end also meant that operations in calculus that we now perform through routine rules would have had to be carried out in their discrete versions. An instructive illustration is the

rearrangement of the second term in the recursion relation for S_n^k as a double sum. It is easy to see that it remains valid when i^k is replaced by any (say real) numbers f_i (the enumerative proof continues to hold; the same identity for another choice of f_i is used in the calculus of the sine series, see Chapter 12.2 below)[3]:

$$\sum_{i=1}^{n-1}(n-i)f_i = \sum_{i=1}^{n-1}\sum_{j=1}^{i} f_j.$$

In the 'continuum limit', it becomes an identity for integrals involving the real function f:

$$\int_0^w (w-u)f(u)du = \int_0^w du \int_0^u f(v)dv.$$

To prove it in this calculus form, write

$$g(w) = \int_0^w f(u)du$$

so that $f(w) = dg(w)/dw$ (from the fundamental theorem, let us not forget). The left side is then

$$wg(w) - \int_0^w u\frac{dg(u)}{du}du = wg(w) - wg(w) - \int_0^w g(u)du$$

on integration by parts, and it is the right side of the identity. Jyeshthadeva's rearrangement formula is nothing other than discrete integration by parts.

It is an amusing diversion now to copy *Yuktibhāṣā*'s 'first principles' derivation of the integrals of (positive integral) powers in the language and using the tools of modern calculus, i.e., to prove directly, by induction on k, that

$$J_k(t) = \int_0^t u^k du = \frac{t^{k+1}}{k+1}$$

for $k = 0, 1, 2, \cdots$. Assume that the formula is true for a fixed k. The *saṃskāram* step for J_{k+1} is to break it up as a first guess and a correction:

$$J_{k+1}(t) = tJ_k(t) - \int_0^t (t-u)u^k du.$$

The first term is $t^{k+2}/(k+1)$ by the induction ansatz. On integrating by parts as above, the second term becomes

$$\int_0^t (t-u)u^k du = \int_0^t du \int_0^u v^k dv = \frac{1}{k+1}\int_0^t u^{k+1}du = \frac{1}{k+1}J_{k+1}(t).$$

[3]Readers will recognise it as a special case of what is known as the Abel resummation formula.

So

$$J_{k+1}(t)\left(1+\frac{1}{k+1}\right)=\frac{t^{k+2}}{k+1}$$

or

$$J_{k+1}(t)=\frac{t^{k+2}}{k+2}$$

which was to be proved. This is not the kind of proof normally found in textbooks. Apart from its unexpectedness and elegance, it has the merit for us of making transparent, in a language we are at home with, how the *saṃskāram* ansatz of replacing one power of the variable by the upper limit of integration leads almost inevitably to an inductive procedure. It also separates the integral aspects of the problem from the differential; no prior knowledge of the derivatives of powers is required.

11.5 The Arctangent Series

Yuktibhāṣā takes up the general arctangent series, θ as a power series in $\tan\theta$, $0<\theta\leq\pi/4$, after the basic π series is established. It is a characteristic of Indian mathematical methodology that one often starts with a special case which is then generalised, sometimes in several stages (and that one does not hide it in the exposition). We have met quite a few examples of this preference, the preeminent one being of course the diagonal theorem: from the square, for which it is more or less self-evident, to the rectangle. Recall also Brahmagupta's approach to cyclic quadrilaterals: rectangles to pre-Brahmagupta quadrilaterals to Brahmagupta quadrilaterals to general cyclic quadrilaterals (and, related to it, the progression from integral diagonal triples to more general integral geometry). Here again, the difficulty of proving – or even formulating – the two main theorems of Brahmagupta increases progressively with increasing generality (see Chaper 8.4-5).

The generalisation of the π series to the arctangent series fits well into this particular-to-general pattern. But, unlike in the examples cited above, the special case has already been seen to be demanding to visualise and prove. The general case is no harder. In the textbook accounts exemplified by section 1 of this chapter, all we have to do is to change the upper limit of the definite integral from $t=1$ to $t=\tan\theta$. The generalisation is equally straightforward in *Yuktibhāṣā* and is of no technical novelty in itself, though it is not presented in quite the same way. The interest mostly lies in what we can read into the accompanying explanatory comments (and, it may be added, in explanations which could have been given but are absent).

But, first, the proof itself, very briefly. Given a point P on the first octant of the circle, let the extended radius through P meet the side AB of the enclosing square at X (as in Figure 11.1). If θ is the angle subtended by the arc AP at O, then AX is $\tan\theta$ (for unit radius). (*Yuktibhāṣā* takes the trouble of equating $AX=AX/OA$ to the ratio of the half-chord ($\sin\theta$) and the projection

of OP on OA ($\cos\theta$) by similar triangles so as to bring in $\sin\theta$ and $\cos\theta$ as conventionally defined on the circle). Now divide AX into n equal parts each of length $\delta t = \tan\theta/n$. All the subsequent steps are the same as in the series for $\pi/4$ except that we have to substitute $\delta t = 1/n$ by $\delta t = \tan\theta/n$, resulting in the series (writing it out again)

$$\theta = \tan\theta - \frac{\tan^3\theta}{3} + \frac{\tan^5\theta}{5} - \cdots .$$

The process of determining θ as a function of $t = \tan\theta$ is given a technical name; the name is *cāpīkaraṇam* which translates literally though inelegantly as 'arcification', leaving no room for doubt that the problem was seen as one of rectification. And, for the first time, the term 'function' can be employed accurately here, in its original sense of the explicit dependence of a (real) quantity θ on the (real) variable t. Through this one step, the metaphysics has changed from merely making a finite table, however finely spaced, to finding a rule in the form of a functional expression valid for every value of the variable. What are the possible values of t and θ? Nilakantha's stated conviction that π, defined as the ratio of the circumference to the diameter, is an irrational number takes on a deeper meaning in this setting. It would no longer have been possible to exclude numbers that are not rational from geometry; a line or an arc could not be thought to consist of rational points alone since a rational point on the side of the enclosing square determines a point on the arc that cannot be rational and *vice versa*, because the conversion involves π. In brief, the Nila results on the geometry of the circle logically require the notion of the real line to accommodate them even if the limiting operation itself is formulated entirely in terms of reciprocals of natural numbers.

To interpret the Nila mathematicians' pioneering work in terms of functions on the real line or on the circle may appear temptingly natural from our present viewpoint but *Yuktibhāṣā* does not dwell on such questions. (Nilakantha might have, if he had got around to it). It does not even, surprisingly, refer to Nilakantha's irrationality conjectures. What it does is to specialise the arctangent series to another rational multiple of π for θ, namely $\theta = \pi/6$, $t = 1/\sqrt{3}$:

$$\frac{\pi}{6} = \frac{1}{\sqrt{3}}\left(1 - \frac{1}{3\cdot 3} + \frac{1}{5\cdot 3^2} - \cdots\right),$$

except that Jyeshthadeva does not present this as a simple specialisation of the general arctangent series, done with in one sentence such as "choose the arc to be one-twelfth of the circumference". After announcing that "it is in conformity with the particular line of reasoning [of the arctangent series]", he in effect gives an independent derivation, modelled on the $\theta = \pi/4$ case, complete with the geometry of the regular hexagon – yet another instance of the reluctance to exploit the power of general theorems. He does not also make the point that the ratio of consecutive terms has gained a factor of 1/3 over the basic π series, making it that much more efficient in numerical computation (and also, incidentally, absolutely convergent). Nor that one can do even better by

specialising to suitable smaller multiples of π, e.g., $\pi/10$ or $\pi/12$. It is reasonable to conclude that the arctangent series was studied for its intrinsic interest and not with numerical use in mind.

But Jyeshthadeva does worry about convergence. The introductory sentence in the section on the arctangent series is: "Here [I] describe how to 'arcify' the sine or the cosine, whichever is the smaller. Suppose first that the sine is the smaller. ..." The tangent never having been considered as anything other than derived from the sine and the cosine, the meaning here is that the powers of the sine will be the numerators in the first octant; the powers of the cosine in the denominators occur naturally in an expansion in powers of the sine. Further on, the co-arc (*koṭicāpam*) is defined as the complement of the arc, followed by: "If the co-arc (equivalently, the cosine) is the smaller, first make the co-arc", i.e., determine the arc from its cosine (expand in powers of the cosine). In other words, choose the right variable depending on whether the arc extends into the second octant or not; the expansion is to be in that variable, $\tan\theta$ or $\cot\theta$, which is less than 1 in magnitude. The cotangent expansion is not given, presumably because it can be worked out along the lines of the tangent expansion.

Apart from describing in adequate technical detail the contents of Chapter 6 of *Yuktibhāṣā*, my aim in this chapter has been to bring out what it is in them that was new – utterly and unquestionably new – to the world of mathematics of the time. From among the many riches to be found there, pride of place must be given to the conceptualisation of a programme for handling geometrical problems involving curvature and its execution in the rectification of arcs of circles. Independently of my interpretative comments, it is hoped that no doubt remains that the concepts are the founding principles of the discipline we now know as calculus, adapted to the problem at hand but envisioned in a manner general enough to serve as the template for its future evolution. The hallmarks of the subject – the notion of a function and of limits, the approach to the infinitesimal, the basic operations of linearisation (differentiation) and integration and the relatioship between the two – are all unambiguously present. As for the techniques used to implement the concepts some, such as the resort to infinite series expansions, are the same as in 17th century Europe but many are singularly original to India, without a direct parallel in later European work, of which the extensive reliance on recursive reasoning is the most characteristic.

Chapter 6 of *Yuktibhāṣā* ends with an account of methods for estimating the remainder of the (very slowly converging) basic π series when it is cut off after an arbitrary finite number of terms and for rewriting the series itself for faster convergence. There is no calculus here; it is all algebraic and it is exciting and exasperating in equal parts. The excitement is partly in the results which are, quite simply, astonishing in their modernity and partly in the preparatory material which is a meticulously explained primer on the abstract algebra of polynomials and rational functions. The frustration comes from the very atypical absence of any *yukti* – in defiance of the very title of the book – for the results cited. I will come to it in a later chapter, after we are done with the sine and cosine series and the surface area and volume of the sphere.

12

The Sine and Cosine Series

12.1 From Differences to Differentials

Let us begin by recalling the roots of Madhava's power series for the sine (and the cosine; I will not always mention separately the cosine series by name) in Aryabhata's table of sine differences. Apart from the introduction of the sine as a natural and powerful tool in the study of the geometry of the circle, Aryabhata's other insight was that the difference of the sines (cosines) of two angles is proportional to the cosine (sine) of the mean of the two angles. For linear functions the difference is a constant and for general functions it is a measure of their deviation from linearity or, equivalently, from the rule of three. Because of the dual relationship between the sine and the cosine coming from their functional differences, the second difference is proportional to the function itself. This was the key property that Aryabhata used, along with the 'intial values' of the functions for vanishing angles, to carry out an iterative finite integration of the second difference equation for a step size of $\pi/48$ to compile the values of $\sin n\pi/48 - \sin(n-1)\pi/48$ for $n = 1, \cdots, 24$ (Chapter 7.3-4).

Aryabhata's super-compressed verses have no room for any explanation of the thinking behind his formula. But, as we saw in Chapter 7, both the way in which the formula is presented in *Gaṇita* 12 and the only logically sensible reading of the sequence of verses preceding it, in particular *Gaṇita* 11 with its reference to an arbitrary division of the arc, are a strong argument in favour of the above interpretation; it is in fact difficult to find a reasonable alternative interpretation which fits all the circumstances around the stanzas *Gaṇita* 10 - 12. Together with the observation that functional differences add up to the difference in the values of the function at the end points of its domain, what Aryabhata did amounts to setting up the foundation for the discrete calculus of the sine function (see the discussion at the end of Chapter 7). Indeed, on summing up again the first differences, Aryabhata's table becomes a finite series for the sine of an angle as a sum of the sines of multiples of any fixed fraction of the angle; this, as argued in Chapter 7.5, seems to be how Bhaskara II derived

P. P. Divakaran, *The Mathematics of India*, Sources and Studies in the History of Mathematics and Physical Sciences, https://doi.org/10.1007/978-981-13-1774-3_12

his exact formulae for the surface area and volume of the sphere, though he makes no mention of it either in the *Siddhāntaśiromaṇi* or in his explanatory notes on it.

The nine centuries between Aryabhata and Madhava saw no fundamental progress in the investigation of the basic trigonometric functions; they were, rather, the centuries in which the fallacy of Bhaskara I held sway. Sine tables were made routinely and in fair numbers but they differed from Aryabhata's, if at all, only in minor details. It was as though the very definition of trigonometric functions had got frozen as their values at a finite set of equally separated angles – the table was the function. The two most perceptive geometers of the period, from either end of it, were absorbed in the theory of cyclic quadrilaterals and, of them, Narayana seems not to have taken any interest in trigonometry. The other, Brahmagupta, presents a bit of a puzzle; his astronomical work could not have avoided trigonometry and he did make a table, in fact two. In the *Brāhmasphuṭasiddhānta* the step size is the canonical $\pi/48$ (but the value taken for π in computing R_0 is $\sqrt{10}$) and in the later *Khaṇḍakhādyaka* it is $\pi/12$. But the latter is supplemented by an interpolation formula for $\sin(n\pi/12+\delta)$ valid for any $\delta < \pi/12$ and that is interesting because it would seem to indicate that Brahmagupta disagreed with the dictum of Bhaskara I (or, perhaps less likely, that he was ignorant of it – the two were after all contemporaries and near neighbours). More than its numerical utility, the value of the interpolation formula in the context of Madhava's own work is theoretical; it points to the likelihood that Brahmagupta thought of the sine as a function defined on all of the circle, for every (real) value of the angle.

This was the situation at the time of Madhava's entry on the scene: no new insights after a long and sterile preoccupation with making tables. There is no doubt at all that for him the sine and the cosine were functions on the circle in the modern sense, defined at every point of their natural domain, $0 \leq \theta \leq \pi/2$ to begin with and extended to the whole circle by symmetry. (A similar statement holds for θ as a function of $t = \tan\theta$ over $0 \leq t \leq 1$). The careful description of the basic trigonometric functions at the start of Chapter 7 of *Yuktibhāṣā* makes it abundantly clear that they were understood to apply to every arc within their domain of definition. So does the short account of Madhava's interpolation formula in which – and elsewhere in the beginning of Chapter 7 – the term *iṣṭacāpam* (the desired arc) is used to denote any arc whatsoever. Without that general understanding, the calculus that led to the power series would not of course have made any sense.

We should also note again that, as Aryabhata's value of π (*Gaṇita* 10) comes ahead of his formula for the sine difference (*Gaṇita* 12), so does the π series in *Yuktibhāṣā* precede the sine series, and for the same reason. The sine table involves the conversion factor from radians to minutes of arc, R_0, whose value depends on π: $R_0 = 21600/(2\pi) = 3438$ minutes for $\pi = 3.1416$ (Chapter 7.3). There would not have been much point in Madhava deriving an exact functional expression for the sine as an infinite series in powers of $R_0\theta$ if R_0 (equivalently π) is itself not known exactly (even if as an infinite series). These

are, once again, fundamentally conceptual, not algorithmic or computational, issues; the game is no longer about better and better numerical values in finer and finer angular steps, but about in-principle exact (what we would today call analytic) characterisations of functions. It is a token of Madhava's logical (constructivist?) acuity that he did not simply take π as given but unknown and go ahead and it is an interesting historical insight that the calculus of the π series was (had to be) done before the calculus of the sine series.

The basic philosophy of the calculus of the sine series is the same as of the π series (it could not have been otherwise): to compute the variation in the function sin for an infinitesimal variation of its argument θ, begin by dividing the domain of the argument, namely an arbitrary arc, by a finite number, eventually to be made unboundedly large. What is required is the difference in the sines of two consecutive points of the division (the endpoints of an arc segment) and that was known (probably, already to Aryabhata) to be linear in the cosine of the midpoint of the segment. So, consider an arc of the unit circle subtending an angle θ at the centre and cut it ino $2n$ equal parts from the outset (so as to avoid bringing in, as Jyeshthadeva does, the midpoints of the segments). With $\delta\theta = \theta/(2n)$, the sine and the cosine at the ith point of the division are $s_i = \sin i\delta\theta$ and $c_i = \cos i\delta\theta$, $i = 1, \cdots, 2n$. In line with the $2n$-fold division I have adopted, define the differences as $\delta f_i = f_{i+1} - f_{i-1}$, $f = s$ or c, and specialise the general difference formulae of Chapter 7.3-4 accordingly:

$$\delta s_i = s_{i+1} - s_{i-1} = 2s_1 c_i,$$

$$\delta c_i = c_{i+1} - c_{i-1} = -2s_1 s_i,$$

for $i = 1, \cdots, 2n-1$, with the geometrically obvious initial values $s_0 = 0$ and $c_0 = 1$. These are exact relations valid for any n and they capture the essential geometry of the problem, just as the relation $\sin\delta\theta_i = 1/(nd_i d_{i-1})$ does in the π series. The variations δf_i depend on n through $s_1 = \sin\theta/(2n)$ and, as expected, tend to 0 as $n \to \infty$.

Yuktibhāṣā takes a great deal of time and care, even more than in the account of the π series, to explain the limiting process. The starting point is a derivation of the difference formulae for the canonical division of the quadrant into 24 parts – effectively 48 on account of the midpoints of the segments having to be considered. But their validity for division into arbitrarily small equal parts is acknowledged and underlined for the first time: "Here the arc segment has to be imagined to be as small as one wants. Then the first sine difference will be the same as the arc segment. (In our notation and language, $\delta s_1 = 2\sin\theta/(2n)\cos\theta/(2n) = \sin\theta/n \to \theta/n$ as $n \to \infty$). . . . But since one has to explain in a certain (definite) manner, I have said that a quadrant has twentyfour chords" (7.5.2). It then goes on to ask that each segment be divided by a very large number, exemplified by *parārddham* (10^{17}), exactly as in the derivation of the π series, making it clear, again, that the divisor is to be made as large as one wishes: "The smaller the segment, the more accurate the half-chord will be". Division by unboundedly large numbers is as much of a presence

in the discussion as in the π series and the smallness of the result of such a division is conveyed by phrases such as "atomic" or "of the nature of atoms" or "extremely atomic" until, finally, Jyeshthadeva hits upon a phrase that we will recognise rightaway, *śūnyaprāyam* ("of the nature of zero", 7.5.3), which is as good a rendering in Sanskrit of 'infinitesimal' as we can hope to have. Not only was a new mathematics being invented but also a new terminology to meet its needs.

The passages on the limiting procedure are among the more difficult parts of *Yuktibhāṣā*. The writing seems to fall short of its habitual clarity and, one might even say, becomes somewhat laboured. Madhava might have seen his way to the idea of the infinitesimal clearly enough, but he left no record of his thoughts. *Yuktibhāṣā* is the first account we have of how the Nila mathematicians travelled that difficult path, in fact the only one; Shankara's long verses do not indulge in the patient explication of difficult concepts. In that light, it is understandable that the writing shows signs of the struggle it must have been to convey these totally original ideas in just words. Added to this was an epistemic barrier. At a superficial level, if Jyeshthadeva wanted to use larger numbers than *parārddham* as divisors, he would have been frustrated since, according to his own (and probably widely shared) nominalistic philosophy, they could not be known as they could not be named – at least they were not named in his book; more seriously, as he wrote in the very first chapter, there are 'unknowable' numbers no matter how high one counts. The compromise he made by using the largest named power of 10 as the divisor and then asking the reader to imagine that the result of the division is made vanishingly small, *śūnyaprāyam*, reflects the erosion of the hold of the old nominalism under the pressure of the demands of Madhava's infinitesimal thinking: to get to the vanishingly small, one had first to come to terms with the infinite and unnamable totality of all numbers.

To return to the mathematics of the limit $n \to \infty$, denote $i\delta\theta$ by ϕ, which is now a general angle in its domain $0 \leq \phi \leq \theta$, and rename $\delta\theta (= \theta/(2n))$ as $\delta\phi$; the difference equation for the sine takes the form

$$\sin(\phi + \delta\phi) - \sin(\phi - \delta\phi) = 2 \sin \delta\phi \cos \phi.$$

As $n \to \infty$, $\sin \delta\phi \to \delta\phi$, the left side is the variation $\delta \sin \phi$ in $\sin \phi$ for a variation $2\delta\phi$ in its argument and we would immediately rewrite the equation as

$$\lim_{\delta\phi \to 0} \frac{\delta \sin \phi}{2\delta\phi} = \frac{d \sin \phi}{d\phi} = \cos \phi$$

along with the corresponding equation for the cosine

$$\frac{d \cos \phi}{d\phi} = - \sin \phi.$$

Yuktibhāṣā does not have the notion of the derivative and so only replaces $\sin \delta\phi$ by $\delta\phi$, leaving the equations in the form

$$\delta \sin \phi = 2\delta\phi \cos \phi, \qquad \delta \cos \phi = -2\delta\phi \sin \phi.$$

It is understood that $\delta\phi \neq 0$, it is *śūnyaprāyam* but not *śūnyam*, and in that sense the objects it works with are differentials.

Upto this point it has been all geometry, as in the determination of $d\theta/dt$ in the π series. The infinitesimalisation has also followed the same path of division by infinity. The difference between the two problems is in the nature of the result the geometry led to. While $d\theta/dt$ is an explicitly known function of t, thereby reducing the integration to a quadrature, that is no longer the case here. $d\sin\theta/d\theta$ is not an explicitly known function of θ; to express $\cos\theta$ as a function of θ (which would have enabled an integration by quadrature) is part of the problem. One can invert the derivative, $d\theta/d\sin\theta = (1-\sin^2\theta)^{-1/2}$, expand the right side as the *samskāram* series for the square root and integrate it term by term (integrals of powers) to get an expansion of θ as a power series in $\sin\theta$ but that – thanks to the legacy of Aryabhata – was not what was required. Interestingly enough, Newton did precisely that when his time came but he did not stop there; he reinverted the series by effectively manhandling the first few terms and guessing from their form ("by observing analogies") the regularity of the coefficients.

The Nila approach to integrating the difference/differential equation for sin and cos is entirely different and very modern. Given a pair of coupled first order differential equations of the type satisfied by sin and cos, we would today convert them into one second order linear homogeneous differential equation for either function by cross substitution and that is exactly what *Yuktibhāṣā* does. Define the second differences (adapted for the $2n$-fold division of the domain)

$$\delta^2 f_i = \delta f_{i+1} - \delta f_{i-1}.$$

Cross substitution leads immediately to the same second difference equation (the discrete harmonic equation, to give it a name) for both sin and cos:

$$\delta^2 f_i = -4s_1^2 f_i;$$

for example,

$$\delta^2 s_i = \delta s_{i+1} - \delta s_{i-1} = 2s_1(c_{i+1} - c_{i-1}) = 2s_1\delta c_i = -4s_1^2 s_i.$$

In the asymptotic limit $n \to \infty$, this becomes the very familiar differential equation

$$\frac{d^2 f(\theta)}{d\theta^2} + f(\theta) = 0$$

for $f = $ sin or cos. What *Yuktibhāṣā* does is to solve the discrete harmonic equation for the correct initial values of sin and cos and then take the asymptotic limit.

The debt the approach owes to Aryabhata's method of computing sine differences for the 24 canonical angles is already evident in the recourse to second differences. The execution of the strategy also derives from Aryabhata's discrete integration (see Chapter 7.4 along with the last part of Chapter 7.5),

but now generalised to an infinitely fine division of the arc (*chindyāt . . . yatheṣṭāni*, *Gaṇita* 11). Naturally, the generalisation is subtle (conceptually) and intricate (technically). Before going into the details of what is probably the most impressive and difficult of the many achievements of the Nila school, it is useful to summarise the strategy one more time: first show that the functions sin and cos solve the harmonic equation and then find the appropriate solutions as power series in the argument.

12.2 Solving the Difference/Differential Equation

To convey the main ideas of *Yuktibhāṣā*'s solution of the discrete harmonic equation and the evaluation of its asymptotic limit, I begin by rephrasing it in the language of classroom calculus, i.e., by starting directly with the differential equation and the fundamental theorem of calculus, but otherwise following the steps in the text. Such an introduction in familiar language will help the reader navigate through the necessarily more involved algebra of the discrete problem. Suppose therefore that all we know about a function $f(\theta)$ is that it satisfies the harmonic differential equation (a prime denotes the derivative)

$$f''(\theta) = -f(\theta).$$

Integrate it formally (the fundamental theorem):

$$f'(\theta) - f'(0) = \int_0^\theta f''(\phi)d\phi = -\int_0^\theta f(\phi)d\phi,$$

and again:

$$f(\theta) - f(0) = \int_0^\theta f'(\phi)d\phi = \int_0^\theta \left(f'(0) - \int_0^\phi f(\chi)d\chi \right) d\phi$$

leading to

$$f(\theta) = f(0) + \theta f'(0) - \int_0^\theta d\phi \int_0^\phi d\chi f(\chi).$$

This procedure is an example of the standard way (19th century) of converting certain types of differential equations into integral equations; the point worth noting in the Nila context is that the only principle used is the fundamental theorem and that its discrete prototype is Aryabhata's treatment of the sine table. Moreover, from the geometry one knows that for $f = \sin$, $f(0) = 0$ and $f'(0) = \cos 0 = 1$ and for $f = \cos$, $f(0) = 1$ and $f'(0) = \sin 0 = 0$; the functions sin and cos thus satisfy the integral equations

$$\sin\theta = \theta - \int_0^\theta d\phi \int_0^\phi \sin\chi d\chi$$

and

$$\cos\theta = 1 - \int_0^\theta d\phi \int_0^\phi \cos\chi d\chi$$

respectively. We could have got the same equations without explicitly introducing the second derivative, by first integrating the first order equations satisfied by sin and cos and cross-substituting, which is the variation preferred by Shankara.

One look at these equations with their linear-with-inhomogeneous-term structure is enough to tell us they are made to order for the exercise of recursive refining. The first guess results from ignoring the second term on the right involving, as it does, the unknown function, i.e., in the sine equation, $\sin_1\theta = \theta$ is the first guess, an approximation that gets better the smaller θ is. On substituting this approximate solution for the sine in the integrand on the right, we get a correction:

$$\sin_2\theta = \theta - \int_0^\theta d\phi \int_0^\phi \chi d\chi = \theta - \frac{1}{2}\int_0^\theta \phi^2 d\phi = \theta - \frac{1}{2\cdot 3}\theta^3.$$

Next, replace $\sin\chi$ in the exact equation by $\sin_2\chi$ to get the third approximation to $\sin\theta$ as

$$\sin_3\theta = \theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!}$$

and so on. Alternatively, replace $\sin\chi$ under the integral by its integral representation:

$$\sin\theta = \theta - \int_0^\theta d\phi_1 \int_0^{\phi_1} d\phi_2 \left(\phi_2 - \int_0^{\phi_2} d\phi_3 \sin\phi_3\right)$$

where I have changed ϕ and χ to ϕ_1 and ϕ_2 for notational uniformity. This process too can be continued indefinitely and the result at each stage of the refining,

$$\sin\theta = \theta - \frac{\theta^3}{3!} + \cdots + (-1)^k \int_0^\theta d\phi_1 \int_0^{\phi_1} d\phi_2 \cdots \int_0^{\phi_{2k-1}} d\phi_{2k} \sin\phi_{2k},$$

is exact, a much more sophisticated example of the type of *saṃskāram* identity Nilakantha illustrated the geometric series with (Chapter 10.4) and used in the derivation of the π series. The two variants of the *saṃskāram* differ in their conceptual basis; the first is an approximation scheme while the second establishes a sequence of integral equations of higher and higher order, all of them satisfied by the sine. *Yuktibhāṣā*'s account oscillates between the two points of view and in practice they both lead to the standard power series,

$$\sin\theta = \theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \cdots,$$

either from indefinitely iterated refining, or simply in the limit $k \to \infty$, respectively.

The integral equation for the cosine can be solved in identical fashion, resulting in the standard series

$$\cos\theta = 1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \cdots .$$

I have described this (not so difficult) material in some detail partly so that it may serve as a faithful guide to Jyeshthadeva's derivation and solution of the corresponding discrete integral equations (and also as a tribute to the attention he pays to every detail) and partly because it is not easy to find a modern textbook that treats these series in this particular way. The basic idea really is as simple as that.

To get back to the discrete problem, recall our notation: an arbitrary angle θ is divided into $2n$ equal angular segments of size $\delta\theta = \theta/(2n)$; for $0 \leq i \leq 2n$, $\sin i\delta\theta$ and $\cos i\delta\theta$ are denoted by s_i and c_i. *Yuktibhāṣā* makes it clear from the start that the limit $n \to \infty$ will be taken eventually and that it is therefore legitimate to replace $s_2 = \sin\theta/n$ by $\delta\theta$. As it happens, the algebra and the combinatorics are independent of this replacement and, hence, of the limit approximation. I have chosen here not to make the replacement, postponing the infinitesimal limit until the solution is fully worked out as a trigonometric identity (it will bring out some interesting differences between the calculus content of the π series and the sine series). Jyeshthadeva surely knew this but it is clear from the many asides that he is interested in $\sin\theta$ for every value of θ, as a function, not in some pretty trigonometric identities obeyed by the sine for finite n. (The significance of this remark will become clear presently).

The conversion of the difference equations to discrete integral equations proceeds exactly in the manner described above: add up the differences (the discrete fundamental theorem) $\delta s_{2n-1} = s_{2n} - s_{2n-2}, \cdots, \delta s_1 = s_2 - s_0 = s_2$ to arrive at

$$s_{2n} = \delta s_{2n-1} + \delta s_{2n-3} + \cdots + \delta s_1.$$

Rewrite in terms of cosines by using the difference equations for the sines:

$$s_{2n} = 2s_1(c_{2n-1} + c_{2n-3} + \cdots + c_1) = s_2 + 2s_1(c_{2n-1} + \cdots + c_3)$$

by virtue of the half-angle formula $s_2 = 2s_1c_1$. But each cosine on the right is a sum of cosine differences:

$$c_{2i-1} = \delta c_{2i-2} + \cdots + \delta c_2 + c_1$$

where the last term c_1 is δc_0 by definition, and each δc can be replaced by the corresponding sine, leading to

$$c_{2i-1} = -2s_1(s_{2i-2} + \cdots + s_2) + c_1$$

and thence to

$$s_{2n} = -4s_1^2[(s_{2n-2} + \cdots + s_2) + (s_{2n-4} + \cdots + s_2) + \cdots + (s_4 + s_2) + s_2] + ns_2.$$

Collecting the coefficients of s_{2i}, we get the formula

$$s_{2n} = ns_2 - 4s_1^2 \sum_{i=1}^{n-1} (n-i)s_{2i}.$$

for the sine of the n-fold multiple of an angle in terms of the sines of lower multiples. For $n = 1$ for example we have the trivial identity $\sin\phi = \sin\phi$ and for $n = 2$ the formula $\sin 2\phi = 2\sin\phi - 4\sin^2(\phi/2)\sin\phi = 2\sin\phi\cos\phi$ for $\phi = \theta/n$ (remembering that we began by dividing θ into $2n$ equal parts). For general n it was, of course, entirely new; even today, it is not a formula routinely taught in classrooms and textbooks.

For Madhava (and Jyeshthadeva), the identity for finite n was itself not of primary interest. The interesting part of the formula was the sum over i in the second term on the right. It has the same structure as the right side of the power sum that occurs in the integration of powers in the π series (Chapter 11.4) with s_{2i} in place of i^k; we can therefore rearrange it (discrete integration by parts or Abel resummation) as before. The result is the discrete integral equation

$$s_{2n} = ns_2 - \sum_{i=1}^{n-1} 2s_1 \sum_{j=1}^{i} 2s_1 s_{2j}$$

where I have separated factors which have natural limits as $n \to \infty$: $2s_1$ becomes the differential $d\theta$ with $ns_2 = \theta$ and s_{2j} is $\sin\phi$ as ϕ ranges over the angles from 0 to θ. The formula itself becomes the integral equation for the sine (with the initial values $\sin 0$ and $\cos 0$ prescribed) written down above. It is convenient to call this representation of the sine the *saṃkalitam* – following *Yuktibhāṣā*, which uses the term (rather than *yogam* for sum) for finite series which will eventually become infinite in the limit – representation.

The formula for s_{2n} in either of the two forms given above determines its value in terms of the values of the lower sines $s_{2i}, i < n$, (and the initial value s_1) linearly and recursively and, as such, is the generalisation to arbitrary n and arbitrary θ of Nilakantha's exact solution for the canonical sine table ($\theta = \pi/4, n = 24$) and, ultimately, of Aryabhata's (approximate) rule. In that narrow sense, the formula is not especially exciting. Its value, particularly in its discrete integral equation form, lies in the fact that it is the first step in proceeding to the limit of infinite n and in the computation of the coefficients in the resulting series as closed combinatorial expressions.

The evaluation of the double sum (the discrete equivalent of solving the integral equation) illustrates the power the method of recursive *saṃskāram* had attained in the hands of the Nila mathematicians. It is initiated in either of two ways (also as in the solution of the integral equation, see above) both leading to the same final answer. In one way of doing it, one makes a first guess by replacing all sines, unknown to begin with, by their arguments (the rule of three), $s_{2i} = i\theta/n$, computes the error as a function of the sines, replaces these sines in the error again by the angles and so on, as in the many examples of

recursive refining that we have already met. Alternatively and more faithfully to Nilakantha's approach to the geometric series (Chapter 10.4), one substitutes s_{2i} in the second term on the right in the *saṃkalitam* representation by its *saṃkalitam* representation:

$$\begin{aligned} s_{2n} &= ns_2 - 4s_1^2 \sum_{i=1}^{n-1} \sum_{j=1}^{i} \left(js_2 - 4s_1^2 \sum_{l=1}^{j-1} \sum_{m=1}^{l} s_{2m} \right) \\ &= ns_2 - 4s_1^2 s_2 \sum_{i=1}^{n-1} \sum_{j=1}^{i} j + (4s_1^2)^2 \sum_{i=1}^{n-1} \sum_{j=1}^{i} \sum_{l=1}^{j-1} \sum_{m=1}^{l} s_{2m}. \end{aligned}$$

The procedure is iterated by replacing s_{2m} in the quadruple sum by its *saṃkalitam* representation and so on. The general term depends on θ as $(-4s_1^2)^k s_2$ with coefficients which are even multiple sums (*saṃkalitasaṃkalitam*, sums of sums . . . of sums) of natural numbers. To express the coefficients in a concise manner, define recursively the quantities

$$S_k(i) = \sum_{j=1}^{i} S_{k-1}(j)$$

with $S_k(0) = 0$ and $S_0(i) = i$. (We have already met the notation in Chapter 7.1, in connection with Aryabhata's *citighana*: $S_1(i) = i(i+1)/2$ and $S_2(i) = i(i+1)(i+2)/6$). Indexing the terms by k with $k = 0$ for the first term, the coefficients of the first two terms are $n = S_0(n)$ and $S_2(n-1)$ ($k = 0$ and 1 respectively). Moreover, we have

$$\sum_{j=1}^{i} S_k(j-1) = \sum_{j=0}^{i-1} S_k(j) = \sum_{j=1}^{i-1} S_k(j) = S_{k+1}(i-1);$$

the quadruple sum which is the coefficient in the third ($k = 2$) term then becomes

$$\sum_{i=1}^{n-1} \sum_{j=1}^{i} \sum_{l=1}^{j-1} \sum_{m=1}^{l} m = \sum_{i=1}^{n-1} \sum_{j=1}^{i} S_2(j-1) = \sum_{i=1}^{n-1} S_3(i-1) = \sum_{i=1}^{n-2} S_3(i) = S_4(n-2)$$

using, once again, the vanishing of $S_k(0)$. The coefficient in the general term is, obviously, $S_{2k}(n-k)$. For a general angle θ, we have thus the trigonometric identity

$$\sin\theta = \sin\frac{\theta}{n} \left(S_0(n) - 4\sin^2\frac{\theta}{2n} S_2(n-1) + 4^2 \sin^4\frac{\theta}{2n} S_4(n-2) - \cdots \right).$$

For any finite n, the series terminates after n terms, since $S_{2k}(n-k)$ vanishes for $k = n$.[1]

[1]The somewhat heavy description above, in the accepted modern way of dealing with the general case once and for all, may put some readers off. *Yuktibhāṣā* of course does everything

This formula for $\sin\theta$ as a polynomial in $\sin\theta/2n$ (apart from the overall factor $\sin\theta/n$) for arbitrary n is exact. Even though the steps in my derivation are the same as in *Yuktibhāṣā* – dressed up in modern notation – the exact formula for finite n will not be found there. That is because Jyeshthadeva has, long before the final formula is arrived at, replaced $2\sin\theta/2n$ and $\sin\theta/n$ by θ/n, with many explanations along the way to the effect that when n is made unboundedly large (the arc segment is made "endlessly atom-like"), the arc segment and its half-chord become equal. It is hard to think that Jyeshthadeva (or Madhava) did not see that the combinatorics of the coefficients go through just as well for the powers of $2\sin\theta/2n$ as for the powers of θ. We must conclude that the goal from the beginning was analytic, the determination of $\sin\theta$ as a function of θ, rather than a fancy trigonometric identity.

As $n\to\infty$, the general term in the polynomial tends to $(\theta/n)^{2k+1}$ with a coefficient determined by the asymptotic behaviour of $S_{2k}(n-k)$. The asymptotically dominant term in the *saṃkalitasaṃkalitam* was already given by *Yuktibhāṣā*, together with its elementary inductive proof, at the end of the account of the π series (where, incidentally, it really plays no role) as

$$S_k(n) \sim \frac{n^{k+1}}{(k+1)!}.$$

The powers of n coming from $S_{2k}(n-k)$ cancel against the powers of n coming from the 'variable' θ/n term by term and the polynomial itself becomes the power series for the sine in the limit.

Yuktibhāṣā also gives the exact expression for $S_k(n)$ for general n:

$$S_k(n) = \frac{n(n+1)\cdots(n+k)}{1\cdot 2\cdots(k+1)} = \frac{(n+k)!}{(k+1)!(n-1)!}$$

even though it is not required in the problem at hand. More surprisingly, it does not provide even a hint about its *yukti*, one of only two instances in the whole book of a significant result being cited without proof. (The other concerns a sequence of estimates for the truncation error in the π series, to be dealt with in the next chapter). One other Nila text, Shankara's *Kriyākramakarī*, quotes the formula without, again, a proof but with the comment that he refrains from supplying a proof because it is not easy to follow, not *sugama*. Perhaps the comment only means that no one really was comfortable with combinatorial mathematics, a topic in which the Nila school showed little interest for its own sake.

The formula actually goes back a long way, but in the context of prosody. $S_k(n)$ is the binomial coefficient ${}^{n+k}C_{k+1} (= {}^{n+k}C_{n-1})$ and is therefore the

---

descriptively by enumerating the individual contributions to the first two terms to determine their structure and then states the form of the general term. I have also taken some liberties with the order in which the various parts of the argumentation are put together, see further down. For a more literally faithful account of the material, see P. P. Divakaran, "Notes on *Yuktibhāṣā*: Recursive Methods in Indian Mathematics" in [SHIM].

answer to Pingala's question 5 (Chapter 5.5) on the number of distinct metres (denoted by $N_{n,m}$ in that chapter) of syllabic length n with m *laghu* (or *guru*) syllables when n is identified with $n+k$ and m with $k+1$. In fact the recursive definition of $S_k(n)$ implies immediately the recursion relation

$$S_k(n) = \sum_{i=1}^{n-1} S_{k-1}(i) + S_{k-1}(n) = S_k(n-1) + S_{k-1}(n)$$

which is precisely the recursion relation satisfied by Pingala's $N_{n,m}$ for the appropriate n and m and leads, following Shah's reading of Pingala (and Bharata; see, again, Chapter 5.5), to a construction of the binomial coefficients. The formula appears in its sum-of-sums avatar in Narayana's *Gaṇitakaumudī* but without proof. As noted earlier, Narayana was a master of combinatorics; it is not unreasonable to suppose that he knew his *Chandaḥsūtra* and that he arrived at the formula by exploiting the recursion relations obeyed by $S_k(n)$ by starting with small k (and, possibly, small n) and building up from there (as indeed Pingala himself might have in the prosody problem). The Nila mathematicians were intimately familiar with Narayana's geometry (Chapter 8.5); there is no reason to suppose that his combinatorial results would have been unknown to them, even if none of them worked on such questions.

In any case, to complete the picture, here is an inductive proof in the modern streamlined style, based primarily on the recursive structure. Defining

$$T_k(n) = \frac{(n+k)!}{(k+1)!(n-1)!}$$

we see rightaway that it obeys the same recursion relation as $S_k(n)$:

$$T_k(n) = T_k(n-1) + T_{k-1}(n).$$

So the equality of $S_k(n)$ and $T_k(n)$ will follow if $T_{k-1}(n) = S_{k-1}(n)$ and $T_k(n-1) = S_k(n-1)$. By induction on k and n it is now enough to prove that $T_1(n) = S_1(n)$ for all n and $T_k(1) = S_k(1)$ for all k, which is a matter of easy verification.

Less abstractly, one can start from $S_1(n)$ and $S_k(1)$ and 'induce' from there which, on evidence, certainly would not have been beyond *Yuktibhāṣā*.

12.3 The Sphere

Yuktibhāṣā's derivations of the area A and volume V of a sphere of diameter d do not depend on power series expansions of trigonometric functions. Since they involve the integration of the sine function and its square, they do, however, use the insights gained in the course of working out the series; the material comes after the series are dealt with, in fact at the very end of (Part I of) the book. For this reason and also because the treatment runs almost exactly parallel to

how these questions are still addressed in elementary calculus, it is a fitting topic to conclude this chapter with.

As we saw in Chapter 7.5, Bhaskara II had the correct formulae for both the area and the volume, though how precisely he arrived at them remains something of a mystery. Subsequently, Narayana also gave these formulae but, astonishingly (in mid-15th century!), with the value 3 for π, $A = 3d^2$, $V = d^3/2$, again without explanation. The starting point in *Yuktibhāṣā* is the same as for Bhaskara – the division of the surface by latitudes – but the latitudes are now infinitesimally spaced and the treatment after that is complete and rigorous, supplying the infinitesimal input that Bhaskara's probable rationale (see Chapter 7.5) lacked. Rather than just present it in modern notation – the practice I have generally followed – I give below a translation of the section on the area with the necessary annotation. The passage is short enough for this to be a feasible option; more to the point, it affords a glimpse into the gradual evolution of *Yuktibhāṣā*'s infinitesimal thinking from its first carefully detailed introduction in the π series and it gives, in a small sample, an idea of how a narrative prose style is used to convey precise mathematical thought (as well as of the minor problems of interpretation a modern reader may face).

> Now I narrate how, combining two principles already explained, [namely that] from the sum of half-chords can be produced the sum of second differences [of half-chords] and [that], knowing the circumference and the diameter at one place (for one circle) we can apply the rule of three to pass to any [circle], the area of the surface of a sphere will be produced.
>
> A uniformly rounded object is called a sphere (*gōḷam*). Through the middle of such a sphere, imagine [drawn] two circles, one along east-west and the other along south-north. Then imagine circles, one shifted slightly to the south of the east-west circle and the other slightly to the north. Their distances from the east-west circle should be the same for all parts. Consequently, these two will be slightly smaller than the first one. Then, starting from these, imagine slightly smaller and smaller circles, as described above, all of them at equal distance (measured along the south-north circle, as also elsewhere in the passage) one from another, so as to end at the south and north extremities. Their separation along the south-north circle must be seen to be equal. This being so, imagine that the circle-shaped gap between two [successive] circles is cut at one place, removed and straightened. Then, of the circles on the two sides of the gap, the larger one will be the base and the smaller one the face (the opposite side) of a trapezium whose flanks will be the arc segment which is the separation along the south-north circle [of two successive east-west circles]. Now cut out the part outside the altitude [from the point of intersection of the face with one of the flanks to the base], turn it upside down and join it to the opposite flank. The result is

> a rectangle whose length is half the sum of the face and the base and whose width is the altitude. In this way, think of all the gaps as rectangles. Their widths are all equal. The lengths have varying measures. The area is the product of the length and the width. The widths of all being equal, add the lengths of all and multiply by the width. Thus will arise the area of the surface of the sphere.
>
> Now the method to know how many gaps there are and what their lengths and width are. The radii of the [east-west] circles we have imagined above are half-chords of a circle whose radius is the radius of the sphere. Hence, multiplying these half-chords by the circumference of [a great circle of] the sphere and dividing by the radius of the sphere (i.e., multiplying by 2π) will result in [the circumferences of] circles having each half-chord as radius. These will be the lengths of the rectangles if the chords are taken at the midpoints of the gaps. Multiplying the sum of all half-chords [by 2π] will result in the sum of the lengths of all the figures (i.e., the rectangles). Multiply it by the width to get the sum of the areas [of the rectangles]. The gap between any two [successive] circles is given by the arc segment of the circumference of the south-north circle considered above. Its chord is the width.
>
> Next, the method of getting the sum of the [half-]chords. Multiply the square of the radius of the sphere by the sum of the second differences [of the half-chords] and divide by the square of the full chord of the arc segment. The result is the sum of the half-chords. Multiply this by the width of the figure (i.e., the rectangle). This width is the chord of the arc segment. The sum of the second differences is the first chord segment. Because of the extreme smallness [of the arc segment], these (the chord segment and the width) are [both] almost equal to the full chord. These are multipliers and the square of the full chord is the divisor. But multiplication and division are unnecessary (i.e., the square of the chord cancels from the numerator and the denominator). What remains is just the radius. Since we have to get the area for both halves of the sphere, double the radius. Thus multiplying the diameter of the sphere by the circumference [of a great circle] of the sphere will produce the area of the surface of the sphere.

The overall strategy is clear: divide the surface into narrow strips by minutely separated latitudes, circles parallel to the first east-west circle (the equator), the separation being constant (as measured by the arc segments defined by the intersections of the latitudes with a meridian, which are almost equal to their chords); compute the area of the strips by ignoring the effect of curvature and assuming that each of them is a cylinder (the rectangles of the passage); and then add them up in the limit in which the latitudes are infinitesimally separated.

Typically for *Yuktibhāṣā*, the introductory paragraph sets out the two principles on which the derivation is going to be based. The first, that the sum of second differences (the integral of the second derivative) of the sine is proportional to the sum (the integral) of the sines, goes back to Aryabhata and derives from the discrete harmonic equation satisfied by the sine; it has already been comprehensively explained in connection with the power series. The second is a surprise: one would have thought that the proportionality of the circumference and the diameter did not need to be specially highlighted in a text written in the 1520s; perhaps the recapitulation is occasioned by the fact that Jyeshthadeva was now considering, for the first time, an infinite family of latitudes with continuously varying diameter, from the equator to the pole.

The second paragraph, also typically, explains the geometry in graphic detail. What it amounts to, less graphically described, is the following. Draw the equator and a meridian. Divide the arc of the meridian from the equator to, say, the north pole into a very large number n of equal segments, each subtending an angle $\delta\theta = \pi/2n$ at the centre of the circle that is the meridian and draw latitudes through the points of division. The area of the circular strip between two consecutive latitudes is then very nearly equal to the product of the circumference of the latitude circle through the midpoint of the strip (as stated in the next paragraph), which varies with the latitude, and the width, defined by ignoring the curvature and therefore constant. The area of the surface of the northern hemisphere is then very nearly equal to the product of the width with the sum of all the circumferences.

The next paragraph determines the circumference of the circle at colatitude θ as $(C/r) \cdot r \sin\theta$ $(= 2\pi r \sin\theta)$ where C is the circumference of a great circle and r is the radius of the sphere – this is where the proportionality of C and r is used. The sum of the circumferences of the latitude circles is thus $(C/r)r \sum_{i=1}^{n} \sin\theta_i$ with $\theta_i = i\delta\theta = i\pi/2n$ denoting the ith colatitude.

In the final paragraph, the sum of the sines is evaluated by replacing the sines by their second differences (the discrete harmonic equation), then replacing the sum of the second differences by the difference of the first differences (the cosines) at the end points (very close to the equator and the pole along the fixed meridian). The width is taken to be the full chord of the arc segment, $2r \sin\delta\theta/2$ (which for large n is the same as the arc $r\delta\theta$). The final formula for the area of the hemisphere is thus

$$\frac{1}{2}A = \frac{C}{r} \cdot r \frac{\cos 0 - \cos\pi/2}{2\sin\delta\theta/2} \cdot 2r\sin\delta\theta/2 = Cr.$$

Coming after the careful treatment of the limiting operation in the sine series (as well as in the π series) with its numerous explanations of how the limit is to be approached, the striking feature of the whole passage is the almost casual manner in which the various simplifications arising in the limit are handled: the estimation of the dimensions of the infinitesimal strips, the equality of arcs and chords for small arcs, the replacing of the limits $\delta\theta/2$ and $\pi/2 - \delta\theta/2$ in the discrete integration by 0 and $\pi/2$, etc. The number *parārddham* as a proxy

for an infinite divisor does not make an appearance nor does the canonical angle $\pi/48$; quantities are simply said to be 'small', for infinitesimal. It is as if the metaphysics of the infinitesimal has now become fully internalised, routine, without further need felt to show explicitly that quantities negligible in the limit are indeed negligible. The shifting of the limits of integration which, as we saw in Chapter 7.5, may have been the obstacle that stopped Bhaskara II from giving an *upapatti* for the area formula does not even get a mention.

The other notable point is that it was essentially the same proof that European calculus converged on in its turn and it has remained to this day the natural, most elementary, way of deriving the area formula. The geometry is absolutely identical; the analytical part, which we will today present concisely in the form

$$\frac{1}{2}A = \int_0^{\pi/2} 2\pi r \sin\theta r d\theta = -2\pi r^2(\cos\pi/2 - \cos 0) = 2\pi r^2,$$

differs only in how the integration of $\sin\theta$ with respect to θ is handled; instead of integrating directly the first order equation $d\cos\theta/d\theta = -\sin\theta$, *Yuktibhāṣā* prefers to integrate the second order harmonic equation once to write

$$\int \sin\theta d\theta = -\int \frac{d^2\sin\theta}{d\theta^2} d\theta = -\frac{d\sin\theta}{d\theta}$$

between appropriate limits. Perhaps it only goes to show how central a place Aryabhatan thinking, especially the fundamental theorem, had come to occupy; perhaps it was just a matter of taste.

There are some interesting novelties in the derivation of the volume formula, though the general geometric idea is the same, namely the slicing up of the sphere by latitudinal planes. But the latitudes are now not separated equally in angle (the arc segments of the meridian) but by equal segments of the south-north axis of the sphere. Denoting the colatitude by θ as earlier, $r\cos\theta$ is the distance along the axis from the centre of the sphere to the latitude plane at θ; in other words, $\cos\theta$ is the independent variable, ranging from -1 (at the south pole) to 1 (at the north pole) and the planes are separated by $\delta\cos\theta = 2/n$, n eventually made to tend to infinity. For finite n, the volume V of the sphere is approximated by V_n, the sum of the volumes of cylinders with the variable latitude circle (with the half-chord $r\sin\theta$ as the radius) as base and $2r/n$ as constant height

$$V_n = \frac{2r}{n}\sum_i A_i$$

where A_i is the area of the latitude circle at the ith colatitude (which is now defined for equal segments of $\cos\theta$), the limit $n \to \infty$ to be taken at the end.

Yuktibhāṣā first derives the formula for the area of the circle ($= Cr/2 = \pi r^2$) – as much of a surprise as its stress on the proportionality of the circumference and the diameter in the section on the area formula – by the infinitesimal

cut-and-paste method described in Chapter 7.5, probably due to Aryabhata. The volume is thereby reduced to a sum of the squared half-chords,

$$V_n = r\delta \cos\theta \pi r^2 \sum_i \sin^2\theta_i.$$

The square of the half-chord is the product of the two segments (the *śara*) into which the chord divides the diameter (the south-north axis of the sphere in the present instance). This, as *Yuktibhāṣā* recalls, is an old result of Aryabhata (*Gaṇita* 17), as we have noted in Chapter 7.1; and, as also noted there, it is equivalent to the trigonometric identity $\sin^2\theta = 1 - \cos^2\theta$ for the colatitude. We have therefore

$$V_n = \pi r^3 \delta \cos\theta \sum_i (1 - \cos^2\theta_i).$$

The advantage of trading sines for cosines is that $\cos\theta_i = i\cos\pi/n$ from the equal division of the axis, reducing the sum to the sum of squares of natural numbers (which of course was known from the π series). We can economise on the duplication of some steps by passing directly to the limit:

$$V = \lim_{n\to\infty} V_n = \pi r^3 \int_{-1}^{1} (1 - \cos^2\theta) d\cos\theta = \pi r^3 \left(2 - \frac{2}{3}\right) = \frac{4}{3}\pi r^3.$$

As is its custom, *Yuktibhāṣā* expresses this in terms of circumference and diameter:

$$V = \frac{Cd^2}{6}.$$

Jyeshthadeva is fully aware of the reason for the equal segmentation of $\cos\theta$ in the volume problem in contrast to the equal segmentation of θ in the area problem. After asking the reader to slice the sphere by latitudes as in the area problem, he says: "There (i.e., for the area), the instruction is that the segments of the circumference of the south-north circle should be equal. Here (for the volume), the instruction is that the thickness of the slices should be equal". It is this piece of astuteness – the correct choice of the independent variable according to the problem at hand – that reduces the integration to a quadrature of a power, familiar from earlier work.

These area and volume computations are, technically, far less challenging than the derivations of the π series and the sine/cosine series. Nevertheless, Jyeshthadeva leaves not an atom of doubt that he looked at all these questions from the same infinitesimal perspective; in the volume problem for instance, there is a renewed insistence on the infinitesimal ('atomic') nature of the segmentation of the axis, through the words *aṇuprāyam* and *aṇucchedam*. It is an interesting fact that, in spite of this thematic unity, no Nila text (to my knowledge) attributes these results to Madhava; *Yuktibhāṣā* is the first to even mention them – Nilakantha's *Tantrasaṃgraha* does not. If it is indeed the case that they are late blooms in the garden Madhava planted, we may conclude that originality and innovation, though not of the same profound and transfor-

mative character as in his own work, continued to flourish on the banks of the Nila well after he was gone.

12.4 The Calculus Debates

Those readers who have kept up with the last chapter and the present one would need no further persuasion, hopefully, that they bear witness to a special moment in the story of mathematics, the birth and first steps – increasingly assured by the time we get to the sphere formulae – of the discipline of calculus. The fact however is that historians are divided on what exactly Madhava's programme represents – calculus or not? – in the grand scheme of the evolution of mathematics and it is the purpose of this section to try and understand the basis, mathematical and historiographic, of the continuing ambivalence in certain circles about the true import of the Madhava revolution. A reading of recent writings on the question turns up phrases like "pre-calculus", "a form of calculus", "akin to calculus", and even "ad-hoc limit-increment arguments" (whatever that may mean). More commonly, analyses of the Nila work stop at the trigonometric power series without going into the question of where they originated (and most of them do not even mention the sphere formulae). Some of the points that will come up below we have already met in this book in scattered and incidental contexts; collecting them together in one place, with a sharper focus on the textual evidence, will (again, hopefully) add necessary balance to a debate that is still alive, as well as throw some light on related issues: Was Madhava the first to give the idea of the infinitesimal a well-defined mathematical sense? Were his ideas, methods and results transmitted to Europe after the arrival of the Portuguese on Kerala's shores?

It goes without saying that the benchmark against which to set the calculus of the Nila school has to be the vastly more general form in which European calculus was visualised and formulated by its founders; and that has its own pitfalls. Nowhere is the dilemma of the objective historian of science more acute than here: how can one deliberately unlearn all that one has learned of a discipline that has grown so vigorously and in so many unanticipated directions to where it is today, beyond anything that Newton and Leibniz, leave alone Madhava, might have imagined? In this light, and in the light of some of the later historical writings on the subject, it is quite remarkable that Charles Whish, the first Westerner ever to have come across the Nila corpus, had no doubt about where its true significance lay. Not only does 'Quadrature (more accurately,'rectification', since it is the circumference rather than the area that is the object of study) of the Circle' occur in the title of his lecture, the text itself has many references to fluxions (and, less frequently, to fluents). Here is a quotation:

> Some quotations which I shall make from these three books (*Tantrasaṃgraha*, *Karaṇapaddhati*, *Sadratnamālā*) will shew that a system of fluxions peculiar to their authors alone among *Hindús*, has

> been followed by them in establishing their quadratures of the circle; and a few more verses, which I shall hereafter treat of and explain, will prove, that by the same mode also, the sines, cosines, & c. are found with the greatest accuracy.

Whish adds that the grandson of Putumana Somayaji, the author of *Karaṇapaddhati*, was still alive and that he himself knew Shankara Varman, the author of *Sadratnamālā* ("which abounds with fluxional forms and series"), whom he describes as "SANKARA VARMA, the younger brother of the present *Rájá* of *Cadattanáda* near *Tellicherry*, a very intelligent man and acute mathematician". There is some confusion about *Tantrasaṃgraha* and *Yuktibhāṣā* ("the commentary in the vulgar tongue") in the writing (copies of *Tantrasaṃgraha* often seem to have included material from its commentaries); so, his mention of the author of the *Tantrasaṃgraha* as having " laid the foundations for a complete system of fluxions" is probably a reference to *Yuktidīpikā* or *Yuktibhāṣā*. Whish was the only person in a position to know the European history of calculus to have been in direct contact with an accomplished mathematical descendent of the Nila school. Did he tell his friend how the 'quadrature of the circle' (and the sine and cosine series) was treated by fluxions and fluents where he came from?

The hundred years or so between Whish's lecture and the gradual re-rediscovery of the content of the texts he first brought to general notice – primarily through the writings of C. T. Rajagopal and his collaborators and the book of Sarasvati Amma – were a dark period from the historian's perspective. The new scholarship that followed came with its own sensibilities and preferences. Rajagopal was dazzled by the algebraic and number-theoretic virtuosity exhibited in the various rearrangements of the π series, some of which he connects with results of Euler and Ramanujan, and the related estimates of truncation errors (to be detailed in the next chapter). Sarasvati Amma, given the scope of her book – all of Indian geometry – was largely content to describe the geometric starting points and work out the algebraic intricacies in the derivations of the power series and the sphere formulae, without going too deeply into their shared roots. It is not a surprise then that the few historians who took note of these new revelations, in turn beguiled by the unexpectedness and sophistication of Madhava's discoveries – almost three centuries before Newton and Gregory and Leibniz! – thought of their infinitesimal moorings as just another detail in a difficult (see Sarasvati Amma's book) subject.

The first person after Whish (to the extent of my knowledge) to have brought out the calculus connection explicitly, though possibly somewhat unwittingly, was in fact not a specialist of Indan mathematics but of European calculus: D. T. Whiteside, the editor of Newton's mathematical papers. This came in the form not of learned articles but in annotational footnotes to two of Newton's papers (and, apparently, also in correspondence with Rajagopal; we are in the 1960s by now). The notes are remarks to the effect that the Gregory-Leibniz series had been derived earlier by a "Hindu mathematician" by the same method of '*reductio per divisionem*' (which is remarkably insightful) and

that the modern parallel to one essential step in the Nila derivation of the sine series would be part of what I have described in the beginning of section 2 of this chapter, the recursive solution of the integral equation satisfied by the sine.[2]

Whiteside's insight too seems to have met with indifference, within India and outside. Well after that and independently of it, however, two opposing trends slowly emerged both, ironically, buttressed by the trigonometric power series, neither particularly concerned with where they came from. For the 'pro-calculus' partisans, it is enough that the power series were among the spectacular products of early European calculus; and, for some of them, the question is tied up with a vigorously promoted hypothesis that the series actually travelled from the Kerala coast to Catholic Europe and influenced developments there. For the 'anti-calculus' faction, the fact that European calculus, from the start, concerned itself with several aspects (mostly 'local', to do with derivatives alone, like drawing tangents to curves and associated problems such as the determination of maxima and minima) which are absent in the Nila work is enough proof: the series may be "brilliant", "ingenious" and so on, but did they come out of calculus?

The Nila lack of emphasis on local properties has a very good historical reason. As is evident in Aryabhata's introduction of differences and second differences, and as *Yuktibhāṣā* makes explicit, the path to calculus began in India as a method of quantifying deviations from the rule of three in the simplest geometric situations involving curvature, namely the circle. Drawing the tangent to a circle at a given point is a trivial matter; it is just the normal to the radius at that point, a problem solved already in the *Śulbasūtra* if not earlier. More generally on the question of local properties, one might say that the primary calculus notion in India was the integral, whose exact evaluation then entailed the study of its infinitesimal constituent parts as defined by the simple expedient of division by infinity: the logical progression is from the whole to the atom (*aṇu*) as it were. Thus, while the rule for integration by parts was discovered and exploited, its infinitesimal counterpart, the Leibniz rule $d(fg) = fdg + gdf$, is never mentioned. In Europe, the local and the global aspects were often looked at independently to start with (Leibniz is a partial exception and, to that extent, closer to the Nila viewpoint) and then brought together in the fundamental theorem. In Newton's early papers,[3] it is quite common to find long

[2]The notes are to "Methods of series and fluxions" in [Ne-W] vol. III and "On analysis by equations unlimited in the number of their terms" in [Ne-W] vol. II. Whiteside mentions *Yuktibhāṣā* as well as a review of the Tampuran-Ayyar Malayalam edition of it in *Mathematical Reviews* (by Rajagopal and one of his many collaborators). He goes on to say ("On analysis by equations . . ."): "The Hindu approach, which depends, for example, on repeated iterations of the identity $\sin x = \int[1 - \int \sin x dx]dx$ is ... wholly distinct from Newton's ...". Much later, and somewhat inconsistently, Whiteside explicitly rejected the idea that this actually was calculus ("ad-hoc . . . arguments", see above). His dates for the texts as well as his attribution of their authorship are entirely unreliable; Madhava is not mentioned.

[3]See especially the first three volumes of the collected papers ([Ne-W]) as well as "The quadrature of curves" in vol. VII (recalling earlier work).

passages discussing differentiation and integration as more or less autonomous operations with their own rules. The use of division by infinity as the bridge between the finite and the infinitesimal could not even have been envisaged, given the newness of enumeration by decimal numbers ("the doctrine recently established", see the quotation from Newton in Chapter 10.4 above) and the consequent fear of infinity. Much of the disputation, mathematical and philosophical, that followed the introduction of infinitesimals is to be traced to this dilemma – there really is no way to define derivatives or differentials without bringing in infinite sequences and their limits which, of course, was how the issue was eventually settled.

The origins of the European focus on local properties are also well understood. Firstly, the Greek geometry of conics supplied a class of simple plane curves whose local aspects, in particular tangents, were not so elementary as for the circle. Far greater was the impact of Cartesian geometry: the tremendous advantage for calculus of associating any reasonably good function of a (real) variable with a curve in the plane, its graph, was immediately grasped. In particular, the problem of finding tangents to a curve became that of computing the derivative of an anlytically expressed function which, in turn, was reduced to a set of simple rules: "The calculus becomes an algorithm", as Whiteside titles one of Newton's manuscripts. The facility with which Newton moves between functions and curves and the enormous variety of both that he examines in dozens if not hundreds of illustrative examples are in themselves a tribute to the geometric and algebraic richness that European calculus was born with, as well as to its universal applicability. From there to the intuition that the operations of differentiation and integration are inverses of each other, the fundamental theorem of calculus for functions of one real variable, was but one small step.

Nothing remotely so general had happened on the banks of the Nila. No Nila mathematician drew the graph of the sine function, and doing calculus was not reduced to a set recipes which did not need the genius of a Madhava or a Newton to implement; calculus did not have the time to become an algorithm. What comes through in the solution of the handful of problems the pioneers took up, all arising naturally in the geometry of the circle and the sphere, is the conceptual coherence and technical uniformity they share, even if their great generality is not explicitly acknowledged in the texts. There are deviations from what has become our normative idea of calculus, the way it was to be done in Europe, but they are understandable in a nascent discipline and have natural explanations. To take the most prominent of them, the fundamental theorem did not outgrow its discrete form because it did not need to, as long as the infinitesimal limit is taken after the *saṃkalitam* is completed. Its role is effectively taken over by the careful estimation of neglected terms and the demostration that their contribution to the sum does indeed vanish in the limit. It is nevertheless of indispensable value in turning the discrete harmonic equation into its integral counterpart, a procedure whose originality *Yuktibhāṣā* acknowledges by a careful account of the principle involved and,

rarest of rarities, by a worked out example, the 8th sine in the canonical 24-fold division of the quadrant (followed by the disclaimer that 24 is just a generic number). There are other technical steps whose significance in calculus becomes fully apparent only after the limit is taken, such as the discrete integration by parts. All in all we can say that, at the fundamental level, the Nila approach differs from the path followed in Europe in that it first constructed a parallel discrete calculus and then passed to the infinitesimal limit. But pass to the limit it did; there is no doubt at all that that was always the goal.

The question thus remains: on what grounds are the doubts about the calculus credentials of Madhava's infinitesimal geometry based? It would seem that there are mainly two. One, easier to deal with, is that the great majority of interested specialists were excluded until very recently from a first hand acquaintance with the one true and indispensable source, *Yuktibhāṣā*, by the language barrier. By the time K. V. Sarma's English translation appeared, positions had already been taken, influenced mainly by the Sanskrit (verse) writings of Shankara. For all his talent and diligence, Shankara was not so sensitive to Madhava's conceptual boldness and originality as he was to the technical brilliance; he would not have, otherwise, spent time trying to push the interpolation formula into a role it was not meant for, as though it has the same fundamental significance as the sine series (see Chapter 10.3). This led to all kinds of misunderstandings, among them the notion that the infinitely extended interpolation series was the Nila version of the Taylor (a "Taylor-like") series for the sine. That it is not, and there is nothing further to be added to the discussion in Chapter 10.3-4.

The other ground for scepticism is more substantive: the absence of a clearly stated abstract point of view and, following from it, the attachment to working things out problem by problem. The mathematically informed reader today will, almost reflexively, tend to reject the notion that five individually solved problems, without explicit reference to the principles that unify them, are enough to mark the birth of a new discipline. On the other hand, the reluctance to articulate the structural principles underpinning the chosen approach to a problem or a set of related problems is typical of Indian mathematics as a whole, never mind that its practitioners very likely began their academic life studying the first structuralist of them all, Panini. Even when there is a discernible unifying thread running through a particular development – examples would be combinatorics according to Pingala (Chapter 5.5) and the 800 year history of the theory of cyclic quadrilaterals (Chapter 8.4-5) – it is something that we have to extract from what is actually in the texts; neither Brahmagupta nor Narayana tells us in so many words why they did what they did.

In the end, we have to accept that none of the texts, not even the meticulously logical *Yuktibhāṣā*, was written with future historians in mind, to be analysed closely for clues about every little detail of what they knew and what they did not. Nothing illustrates the hazards of coming to easy judgements more clearly than a question we have already encountered in the paragraphs above: did the Nila mathematicians know the generalisation of Madhava's

("Maclaurin") series to the true Taylor series for the sine (as opposed to the wrong interpolation series)? No text mentions it. Nevertheless, it is an immediate and perfectly obvious corollary of two of Madhava's own celebrated results, the sine/cosine series themselves and the addition theorems:

$$\sin(\theta+\delta) = \sin\theta\cos\delta+\cos\theta\sin\delta = \sin\theta\left(1-\frac{\delta^2}{2!}+\cdots\right)+\cos\theta\left(\delta-\frac{\delta^3}{3!}+\cdots\right).$$

What do we conclude? That, at the practical level, it was thought to be no more useful than Madhava's original interpolation formula which, as far as computations went, did a reasonably good job with far less theoretical effort (no calculus!)? Or, at the fundamental level, that it was no more interesting than Madhava's original power series?

The real contrast between the histories of Indian and European calculus is in what happened to the two traditions after their founders had passed on. The two hundred years following Madhava saw, at best, a consolidation of what he had initiated – and pretty much completed, going by the words of those who came after. There is no evidence – except perhaps one or two straws in the wind – that his work came to be known outside the small geography of the Nila basin. The two hundred years following Newton and Leibniz were, in contrast, a period of explosive growth in mathematical scholarship and creativity in every part of Europe. Calculus itself was transformed in ways which would have been unrecognisable to its founders, creating offshoots in many new directions and revolutionising what were considered settled disciplines like its own well-spring, geometry. Later, in the middle of the 20th century, it even, finally, broke free of its Cartesian moorings by adapting itself to distributions, analytically defined objects which are more general than functions and cannot be represented geometrically as curves. Through it all, the concepts thought up by Madhava and Newton and Leibniz, the notions of local linearisation and integration, never lost their centrality; the fundamental theorem kept in step with every generalisation, only changing its name eventually to Stokes' theorem in the broad setting of differential geometry.

I leave the last word to a distinguished contemporary mathematician and historian of mathematics: "It seems fair to me to compare [Madhava] to Newton and Leibniz."[4]

Given the thematic commonalities between the calculus of India and Europe, it is natural to wonder whether the European developments were in any way influenced by the discoveries of Madhava and his school. We have had occasion to note – and we will revisit the question one last time in the last chapter of this book – that establishing knowledge transmission across cultures separated by vast distances is uncertain business. Instances – the arrival of Ptolemaic astronomy in northwestern India and the spread of the astronomy and mathematics of Aryabhata and Brahmagupta to Islamic Mesopotamia and China –

[4] David Mumford, Review of [Pl], Notices of the Am. Math. Soc., vol. **57** (2010), p. 385

backed by solid recorded evidence are rare. More often, the case for transmission is a circumstantial one, its plausibility depending on the convergence of whatever strands of indirect evidence can be adduced. Temporal priority of the transmitter is an obvious necessary condition but, equally obviously, not sufficient. Especially when it comes to the universal truths of mathematics, to assert that their discoveries are unique events and that they are then disseminated across space and time through cultural exchanges is arbitrary and ahistorical. Given the right cirumstances, it seems reasonable to accept that more than one gifted mind can find its way to some of these immutable truths; one might even argue that that is what characterises the universality of mathematics.

To make a good case for the possibility that European mathematicians of the 17th century had knowledge of what had been done in Kerala two and a half centuries before, neither their common content nor the chronological order is therefore sufficient ground. The promoters of the transmission hypothesis have pointed, correctly, to the continuous Portuguese presence along the Kerala coast from about the time Nilakantha wrote *Tantrasaṃgraha* and through most of the next century, especially strong around the mathematical villages in the Nila basin, as a potential channel of communication. But no trace of anything mathematical that may have been sent across to Europe has been found so far in any of the likely repositories; there is no direct evidence of transmission as of today.

But one can look for less direct evidence; an avenue that is still open is a comparison of the variations in the way these common themes are handled in the two cultures, the mathematical styles so to speak, mutations in the universal mathematical DNA (to borrow terminolgy from current work in comparative human genetics of culturally distinct populations). This is a field that has hardly been touched so far, though its potential importance to historians should be self-evident. The few points of contrast that we have gone over in this section make it already clear that the Indian and European styles of doing calculus are quite distinct one from the other. To go from the general to the specific, consider Newton's treatment of the sine series (in the tract "On analysis . . ." in which Whiteside's footnote occurs, cited above). The starting point is the same as Madhava's, the observation that the sine and the cosine are each other's derivatives up to a sign. Newton's main technical tool being the binomial series for fractional powers, he inverts the derivative:

$$\frac{d\theta}{d\sin\theta} = (1-\sin^2\theta)^{-1/2} = 1 + \frac{1}{2}\sin^2\theta + \frac{3}{8}\sin^4\theta + \cdots$$

and integrates the powers of $x = \sin\theta$ with respect to x, which of course was known. (Not inverting the derivative would have led to the more difficult integration of powers of the sine with respect to the angle). But this results in a series for θ in powers of $\sin\theta$, which is not what he wants. Here is what he does next ("On Analysis . . .", [Ne-W] vol. II):

> If it is desired to find the sine from the arc given, of the equation $z = x + \frac{1}{6}x^3 + \frac{3}{40}x^5 + \frac{5}{112}x^7 + \cdots$ found above (z is the arc and x its

> sine), I extract the root, which will be
>
> $$x = z - \frac{1}{6}z^3 + \frac{1}{120}z^5 - \frac{1}{5040}z^7 + \frac{1}{362880}z^9 \cdots.$$
>
> Let it be noted here, by the way, that when you know 5 or 6 of those roots you will for the most part be able to prolong them at will by observing analogies.

We can learn much from this little passage about Newton's approach to the details of his calculus – the single-minded focus on his main goal, the willingness to try anything that works ("analogies" would have barely qualified as respectable mathematics even in the 17th century) – but the important fact for us here is how little it resembles Madhava's. One might even say that the only thing the two have in common is the idea of expanding the sine in a power series. Newton acknowledges the inspiration he derived from decimal numbers but no one as far as I know has suggested that infinite series as legitimate mathematical objects travelled from India to Europe. Whiteside has no doubts: his footnote cited above goes on to refer to Whish's lecture and then adds, ". . . there can be no question of any influence upon Newton in his independent rediscovery of the series".

If there was transmission of the basic ideas of calculus from Kerala to Europe, the case for it is yet to be made; and that is putting it extremely cautiously.

We have seen in Chapter 7.5 that the germ of Madhava's calculus for the sine as a function of the angle is to be found in Aryabhata's *sūtra* on the division of the quadrant into as many segments as one pleases. The confusing history of how the idea was misinterpreted – or even rejected – has also been described at length there. There still remains a question: is there any evidence, even if tenuous, for an infinitesimal mode of thought from the period between Aryabhata and Madhava? There have been recent claims (notably in an article of Ramasubramanian and Srinivas in [SHIM]) that a passage in the *Siddhāntaśiromaṇi* of Bhaskara II (together with his own commentary on it, the *Vāsanā*) on the motion (or the speed in the present context; the word is *gati*) of a rapidly moving planet such as the moon suggests that he had grasped the notion of an instantaneous speed ('instantaneous velocity' is the phrase used). The idea is obviously worthy of serious examination, and not just as a precursor, if it can be substantiated, of the Nila developments; it will have signalled a recognition that trigonometric functions of geometric variables are not the only objects which can be studied from an infinitesimal point of view, in particular a foreshadowing of the Newtonian paradigm of dynamics as calculus.

The passage (translations can be found in [Pl] and the paper of Ramasubramanian and Srinivas) actually starts with the statement that the difference between the positions of a planet at the same (solar) time on two consecutive days gives a speed which is accurate over that day. The *Vāsanā* explains what is meant: during this interval of time, the planet is to be taken to move with that speed between the two observed positions; and that can only be if it is the

average speed over a day that is meant. It is not said that there is a time in between at which the planet actually has a speed equal to the average speed, as has been inferred by those who would like to see in the passage an example of the mean value theorem of calculus.

Bhaskara clarifies that the average speed provides only a gross description of the motion and then goes on to consider a more refined characterisation which he calls *tātkālika* and which is the word that has been construed as 'instantaneous'. The word *tatkāla* has the connotation of 'that *kāla*' and *kāla* is a duration, not an instant, of time; in other words, what Bhaskara is after is the speed in a smaller interval of time over which it can be taken to be approximately constant (equal to the mean speed), for those planets like the moon whose motion is rapid and non-uniform (over a day). In fact he says that the speed will change *pratikṣaṇam*, from *kṣaṇa* to *kṣaṇa*; *kṣaṇa* is rendered in common speech as 'moment' or 'instant' (though, as will be clear in a moment, Bhaskara did not think of it as an infinitesimally small duration). It is this motion, supposed uniform over a small subdivision of the day, that is termed *tātkālika* motion. Also, time itself is not the parameter with respect to which change is directly measured; omitting the astronomy, what is done is to consider intervals in the angle (of longitude) θ along the trajectory and the sine difference at the end points of such an interval $\delta\theta$: $\delta \sin\theta = \sin(\theta + \delta\theta) - \sin\theta$. Bhaskara gives a rule for this which is just the sine difference formula for small $\delta\theta$: $\delta \sin\theta = \cos\theta \sin\delta\theta$ and he says unambiguously that $\sin\delta\theta$ is to be taken as 225 minutes. That is the key fact; he is working within the standard Aryabhatan framework of the 96-fold division of the circle, certainly not a division into arbitrarily small arcs or, even less, an infinitesimal division of a time interval. The situation is reminiscent of his struggle with the sphere formulae and may have the same explanation (Chapter 7.5): he was aware of the need for finer and finer division of the arc to accommodate motion that changes *pratikṣaṇam* but lacked the technical apparatus to carry it through. In any case, to use evocative expressions like 'rate of change' or 'instantaneous velocity' with their Newtonian overtones and to rewrite Bhaskara's equations for first order changes over a finite angle as differential equations is to overread what he actually says.

Similar claims have been made about the computation of changes in other related parameters in Nilakantha's refined astronomical model; that should be of interest because by then Madhava had already lived and done his work but Jyeshthadeva had not yet written his definitive account of it in *Yuktibhāṣā*. Nilakantha gives no justification of the formula he writes down. In the absence of explicit statements to the effect that what is being computed is the response of a function to an *infinitesimal* change in its argument (in the manner of *Yuktibhāṣā*) – or at least a word like *śūnyaprāyam* or *aṇuprāyam* – claims that it comes from working with derivatives can never be settled. The obvious reason, simply, is that the *first order* variation of a function for a finite change in its argument *is* the differential in the limit; that is how the derivative is defined and computed: $\delta f(x) = f'(x)\delta x$ – for example $\delta \sin\theta = \cos\theta\delta\theta$ to the first order

– becomes in the limit $df(x)/dx = f'(x)$. Words and explanations matter, not just the equations.

As far as the Indian context is concerned, these claims and counterclaims are probably best taken as marking a natural phase in the coming of age of mathematical-hstorical scholarship: on the one side, a lingering Euro-centric mindset sustained by incomplete knowledge of some key texts and, on the other, as though in reaction, an uncritical overinterpretation of snippets from other equally important texts. It is fair to note, nevertheless, that a certain tendency towards reinterpretations of early discoveries in the light of subsequent developments is not new and not peculiar to India. It is widely accepted for instance, at least from late 19th century onwards, that Archimedes anticipated integral calculus[5] in his method of exhaustion for areas and volumes generated by conics. The key idea on which the general method is modelled is a theorem in Euclid that says that the area A of a circle is proportional to the square of the diameter (Book XII, proposition 2). Applied to this most elementary of examples, the method substitutes the intuitively obvious equivalence of the circle with a regular inscribed (or circumscribing) n-gon 'in the limit of n tending to infinity' by a *reductio ad absurdum* argument involving only a finite n: if A is not πr^2, there will exist, for some finite n, an inscribed n-gon with area greater than A or a circumscribing n-gon with area less than A. As Heath himself emphasises, there is no appeal at any point to the notions of the infinitely large or the infinitesimally small or even of a limit of a sequence (compare the Nila derivation, probably going back to Aryabhata, of the area of the circle – Chapter 7.5, Figure 7.5 – and, especially the meticulous *Yuktibhāṣā* treatment of the area and volume of the sphere, Section 3 above). Whether to see in this method a distant anticipation of calculus – why not give the credit then to Euclid's proposition XII.2? – should probably be thought to be a matter of cultural predisposition or, at best, a question of definition: what is calculus? For all the brilliance Archimedes brought to the solution of these problems, we have to say today that it was not calculus because it had none of the characteristics associated later with it, even if the problems are tailor-made for the application of true calculus methods. The situation is similar to the one discussed earlier in connection with the history of trigonometry: calculus is the natural discipline for the investigation of the effects of curvature (deviations from the rule of three), just as trigonometry is for the geometry of the circle, but the Greeks did not anticipate the essence of either the one or the other.

[5]See T. L. Heath (ed.), *The Works of Archimedes*, Dover Publications Inc, Mineola , NY (2006, reprint), Chapter VII of the Introduction titled "Anticipations by Archimedes of the Integral Calculus".

13

The π Series Revisited: Algebra in Analysis

13.1 The Problem

In this final chapter on the mathematics of the Nila school, I return to a topic which was put aside earlier. The problem which it deals with has its origin in the realisation of the extremely slow convergence of the basic π series, making it effectively useless for the calculation of good approximations to its numerical value. Having recognised the problem, Madhava set about overcoming it in an exercise as remarkable for its originality and effectiveness as for its sense of modernity. Two closely related ways of dealing with the slowness of convergence emerge: i) a sequence of better and better estimates, potentially without end, for the remainder in the basic π series when it is terminated after a finite but arbitrary number of terms, as rational functions of the termination point; and ii) reorganisations of the series, based on these 'error' estimates, to produce new series that converge more rapidly. The modernity is reflected not only in the nature of the results but also in the methods that lead to them: a recourse to algebraic ideas going well beyond what had become standard by the time of Bhaskara II (as expounded in *Bījagaṇita*, for example). That in turn required attention to be given to the structural properties of certain algebraic objects like polynomials and rational functions, properties defined by the operations that can be carried out on them. In other words, we have here a very conscious preoccupation with algebra in the abstract – what is a polynomial in an unknown (*avyaktarāśi*)? – not just arithmetical manipulations in which one or more numbers happen not to be known though fully or partially determinate. A degree of abstraction is present, implicitly, in all algebraic methods, notably in Brahmagupta's *bhāvanā* and its outgrowths, but it is in the philosophy and execution of approximation techniques for the π series that it is given prominence in its own right for the first time.

P. P. Divakaran, *The Mathematics of India*, Sources and Studies in the History of Mathematics and Physical Sciences, https://doi.org/10.1007/978-981-13-1774-3_13

Once again, our guide in following this new direction will be *Yuktibhāṣā*. There is in fact no choice; though Shankara's writings have the statements of the main results, it is only in *Yuktibhāṣā* that we find a description, meticulous and extensive, of the algebraic foundations and of the details of how they lead to the stated results (with one significant omission which will concern us later). Questions are posed in the greatest possible generality and answered likewise. Surprisingly for a problem whose ostensible goal is to determine finer and finer numerical values for π, no computation to gauge the efficacy of the method is offered. There occur in fact no numbers except as coefficients of certain polynomials (whose roots are not required to be determined); it is 'pure mathematics' with a vengeance.

The long last section of Chapter 6 of *Yuktibhāṣā* begins by enunciating the problem: "Now [I] describe how the results from division by higher and higher odd numbers (the sum of terms up to the reciprocals of higher and higher odd numbers) are refined by means of a final correction (the word used is *saṃskāram*) to bring them closer to the circumference". It is made clear immediately that the form of the correction is to be valid for the sum up to an arbitrary number of terms: "First [we] have to assess (find a criterion for) whether this correction is accurate (*sūkṣmam*) and for that [we] have to do *saṃskāram* on the result of (the sum up to) the reciprocal of any odd number". In executing the programme, it turns out to be convenient – one of several such notational conventions – to label the terms in the series not by the positive integers but by the odd numbers which are their denominators. Write $j = 2i-1, i = 1, 2, \cdots$, so that the general term $(-)^{i-1}(2i-1)^{-1}$ in the π series becomes $(-)^{(j-1)/2}j^{-1}$, j running over all odd positive integers. The remainder after the term with j as the denominator is written as $1/r(j)$, defined (paying attention to the alternating signs) by

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \cdots \mp \frac{1}{j-2} \pm \frac{1}{j} \mp \frac{1}{r(j)}.$$

Correspondingly, $1/r(j-2)$ is defined by

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \cdots \mp \frac{1}{j-2} \pm \frac{1}{r(j-2)}$$

and we have the fundamental relation

$$\pm\frac{1}{j} \mp \frac{1}{r(j)} = \pm\frac{1}{r(j-2)}$$

or

$$\frac{1}{r(j-2)} + \frac{1}{r(j)} = \frac{1}{j}$$

holding for all odd j. Clearly,

$$\frac{1}{r(j)} = \frac{1}{j+2} - \frac{1}{j+4} + \cdots$$

is positive for all j. Jyeshthadeva says nowhere that $r(j)$ is irrational; perhaps he was unaware of Nilakantha's conjecture on π, perhaps the question was not of interest to him.

Setting up the apparently trivial identity for the remainders is the key step. The bold idea now is to think of this remainder equation as determining r as a function of j and to solve it by making a first guess for $r(j)$ and improving on it in a series of steps. The idea of exploiting the functional equation is an amazingly modern one; and to describe the process of iterative refining (in principle, *ad infinitum*) by which the equation is solved as *saṃskāram* is perfectly in keeping with the use of the term elsewhere in *Yuktibhāṣā.*

Jyeshthadeva is acutely aware of the functional nature of the problem as well as of its novelty, and explains the issues at some length. After noting that the remainder identity would be satisfied if $r(j-2)$ and $r(j)$ were both put equal to $2j$, he says:

> It can never happen that each of the *saṃskāram* denominators ($r(j-2)$ and $r(j)$) is equal to twice the odd number (j). Here is why. Suppose the first *saṃskāram* denominator corresponding to the last odd number is double the odd number above it (i.e., $r(j-2) = 2j$). Then the second *saṃskāram* denominator has to be double the next odd number (i.e., $r(j) = 2(j+2)$), since it must be stated in the same manner (or, following the same rule). Then it will be four more than double the first odd number. If [alternatively] the higher *saṃskāram* denominator is taken to be twice the odd number, then the lower *saṃskāram* denominator will be four less. There is no way in which both the *saṃskāram* denominators can be equal to twice the odd number (or even just equal to each other, we may add).

These are a lot of words to convey the notion of functional dependence. One gets the clear sense that the concept of an algebraically specified function in a context in which the only equations are algebraic identities holding for all values of the variable (j) and without the visual prop of geometry was not easily arrived at; it is to bring this out that I have quoted the passage in full. We may also note how close the reasoning comes to a primitive form of *reductio ad absurdum* – the hypothesis $r(j-2) = r(j)$ for all j contradicts a nontrivial dependence of r on j, "it can never be . . ." – without quite going over the line.

The tentative choice $r(j-2) = r(j) = 2j$ is not as silly as may appear at first sight if it is thought of as the zeroeth guess in a process of iterative refining: $2j$ is the value to which both $r(j-2)$ and $r(j)$ tend as j tends to infinity. Effectively, this is how *Yuktibhāṣā* approaches the problem of approximating the remainder and the asymptotic behaviour of $r(j)$ is constantly present in this entire section of the book. In fact, one way of looking at the whole discussion is to think of subsequent improvements on this very crude starting point as better and better estimates of the asymptotic behaviour of the remainder, though it is expressed in a slightly (and interestingly) different language, as we shall soon see.

The working principle now is that r should respect the correct functional dependence as j changes, while bringing $r(j-2)$ and $r(j)$ as close to $2j$ as possible. The first step towards this goal (the first guess in the *saṃskāram* sense) is to take

$$r_1(j-2) = 2(j-1)$$

and, hence,

$$r_1(j) = 2(j+1).$$

Then

$$\frac{1}{r_1(j-2)} + \frac{1}{r_1(j)} = \frac{1}{j - \frac{1}{j}} = \frac{1}{j}\left(1 + \frac{1}{j^2} + \cdots\right),$$

showing that $r_1(j)$ satisfies the remainder identity up to terms of order $1/j^3$. The choice of r_1 seems to have been made on the basis of simple inspection and is not the result of a defined *saṃskāram* procedure but a first guess. *Yuktibhāṣā* is somewhat obscure on this point; in any case it works marvellously.

Rather than look at the asymptotic behaviour, *Yuktibhāṣā* esimates the error or grossness (*sthaulyam*) of this first guess by computing the exact difference

$$\frac{1}{r_1(j-2)} + \frac{1}{r_1(j)} - \frac{1}{j} = \frac{1}{2(j-1)} + \frac{1}{2(j+1)} - \frac{1}{j} = \frac{1}{j^3 - j}.$$

This calculation is of course a trivial exercise in algebra but was thought to need an explanation, as it deals in *avyaktarāśi*, in Jyeshthadeva's time. To get the result, as he says, we must reexpress the reciprocals of the "two denominators and the odd number" in terms of a common denominator so as to be able to add the first two terms and subtract the third. That requires a meaning to be given to the addition (including subtraction), multiplication and division of polynomials in j. It is at this point that he, having already quoted Brahmagupta's verse on multiplication among positives and negatives – which, from the context, was meant to apply to signed variables as well – supplies a little tutorial on what polynomials are and how they are to be manipulated algebraically: "[Normally], we can find the common denominator only when we know the numbers [explicitly]. [The procedure for] when the numbers are fixed may not apply in all situations (i.e., how is the procedure to be generalised to variables?). But there is a method for producing common denominators even without knowing the numbers".

But, before that, I explain briefly what *Yuktibhāṣā*'s criterion for the accuracy of approximations to r is and how it is formulated. It is first noted that taking $r = r_1$ leads to a grossness that is positive, $r_1(j-2)^{-1} + r_1(j)^{-1} > j^{-1}$, and a tentative attempt made to reduce it by adding a constant positive quantity to r_1. The trial choice for the constant is 1: replace $r_1(j-2) = 2(j-1)$ by $2j-1$ and $r_1(j) = 2(j+1)$ by $2j+3$. The resulting grossness when reduced to a common denominator has $-2j+3$ (instead of 1) as the numerator, as is easily checked, while the denominator remains a cubic polynomial. The choice is rejected because the grossness has got worse: it is of order j^{-2} rather than j^{-3} as earlier. This observation is the starting point of the search for a better

additive correction. That the grossness or, equivalently, the left side of the remainder equation, $r(j-2)^{-1}+r(j)^{-1}$, should be two degrees higher in j^{-1} at each stage of the refining – the first correction of order j^{-3} will be sought to be improved to one of order j^{-5} at the next stage – becomes a guiding principle in the *saṃskāram*.

13.2 Polynomials: A Primer

It is not as though the algebraic manipulations required in working out these corrections were unknown earlier. To mention a prominent example, Brahmagupta's theory of the quadratic Diophantine problem and its follow-up by several others (Chapter 8.2-3) involve polynomials of low degree with integral or fractional variables and coefficients and operations on them, including multiplication and division. So do some of the less famous Diophantine equations considered by Bhaskara II (of explicit degree up to 8) in the last chapters of *Bījagaṇita*, as well as the algebraic work of Narayana. We must therefore suppose that algebraic computations with polynomials, not restricted to one variable, were well understood in practice by the time of Madhava; they certainly would not have been an obstacle to doing the algebra he needed in estimating the remainder.

So, why does Jyeshthadeva spend so much time and space (four pages in Sarma's edition of *Yuktibhāṣā*) explaining in general terms what polynomials are and how they are to be handled in computations (he concentrates on the multiplications required in finding a common denominator)? One probable answer is that the notion of the degree of a polynomial plays a crucial role in his criterion for the accuracy of successive *saṃskāram*s of the remainder. The more fundamental reason in my view is his concern for foundational issues; a close parallel will be his explanation of how the integers are characterised fundamentally, independently of the base used in practical counting, by their property of succession (see the discussion in Chapter 4.1-2). As there, it would seem that the intention was not merely to instruct students in how to do computaions (numerical there, algebraic here), but to convey an idea of a polynomial (in one variable) – in other words, to define it – as an object in its own right obeying its own operational rules. The definition uses decimal numbers ('polynomials in 10') as a model, thereby completing the link between numbers and polynomials. The picture that emerges is, of course, incomplete and somewhat fuzzy, far from the picture we have today of the ring of polynomials in the abstract, but the first hesitant signs of an abstract approach are clearly discernible. We may also recall, again, how Newton defined power series in analogy with the "doctrine . . . for decimal numbers" (see the quote in Chapter 10.4).

The account begins, as already noted, by recalling Brahmagupta's famous verse on the rule of signs for multiplication of positives and negatives, whether fixed numbers or algebraic symbols. Then it explains what is meant by (positive integral) powers of a variable or unknown (*avyakta rāśi*, generally shortened to

just *rāśi*), by analogy with the correspondence of places in decimal number notation with powers of 10: every higher power of the variable is associated to the next place to the left, beginning with the zeroeth power at the extreme right. It is then explained how to represent the idea in a written symbolic notation, one of the very rare instances of such a device being used in a mainstream text. The places are demarcated by horizontally aligned compartments or boxes (I use vertical strokes here to separate the boxes) which are occupied by the corresponding coefficients. Thus, the general polynomial is written simply as

$$a_n x^n + a_{n-1} x^{n-1} + \cdots a_1 x + a_0 = |a_n|a_{n-1}|\cdots|a_1|a_0|,$$

the variable remaining implicit (j in the problem at hand). The extension to rational algebraic expressions also follows the convention for numbers: just write it as a fraction, with the denominator polynomial under the numerator. It is made completely clear that the polynomials considered are of general degree; an anonymous verse in Sanskrit (isolated and unattributed verses in this section of *Yuktibhāṣā*, some of which also occur in Shankara's writings, are almost certainly due to Madhava[1]) is quoted which says: "[In a polynomial] the places (omitting the first) are the unknown, its square, cube, the square of its square, its fourth, fifth, sixth powers and so on". If a particular power is absent, it is suggested that its coefficient, zero, be indicated by whatever "object" (token) is used to denote zero (though, logically, it is enough just to leave the box empty; I will use 0 here). Thus the constant polynomial a is expressed symbolically as $|a|$ and the variable itself as $|1|0|$.

The ways in which this place-value representation of polynomials and the concomitant notation deviate from the standard decimal representation are also brought out. Firstly, a coefficient in a polynomial can be a negative number. It is recommended that some marker, not specified in the text, be attached to the number symbol to indicate negativity; I will use a bar over the number here: $4x^2 - 4$ is $|4|0|\bar{4}|$ for example. Connected to this, the coefficients in a polynomial are from the set of all integers, not just the atomic numbers 1 to 9 as in decimal numbers where the entries are integers mod 10 (hence the absence of negative entries). In the second refining of the remainder, the polynomial $(4j^5 + 16j)$ occurs in an intermediate step as the common denominator and Jyeshthadeva uses the occasion to explain that the 16 must be kept as it is in the second box from the right even though it is greater than 9:

> The common denominator has six places in six compartments (it is of 5th degree): [there is] zero in the first compartment; sixteen in the second; in the next three, zero; then four in the sixth compartment

[1]Some of them were in the past attributed to Nilakantha, first in [YB-TA], probably because some manuscripts of *Tantrasaṃgraha* included parts of Shankara's commentary *Yuktidīpikā*, which quotes the same verses. We will probably be safe in concluding that Madhava never wrote up his various discoveries as monographs, contenting himself with scattering nuggets of knowledge among his disciples. Nilakantha mentions two unanticipated occasions on which he came across *śloka*s supposedly composed by Madhava (see Chapter 9.4).

> (i.e., the polynomial is $|4|0|0|0|16|0|$). . . . Here, the number in a compartment will not occupy the higher compartment (to the left). Even if it is increased by ten, it can only be made to occupy (be promoted to) a compartment corresponding to (a power of) ten. But since the *rāśi* is an unknown number, there is no way for it (the coefficient) to occupy a place (with respect to) a number equal to the *rāśi*. If all the numbers that are in (that contribute to) one place are positive or all negative, they must be added; if of both signs,they must be subtracted (appropriately). That is all that can be done.

This may strike a modern reader as a clumsy way of distinguishing between the values that a variable takes and the set from which the coefficients are chosen, but there can be no doubt that the distinction was well understood – in modern terminology, Jyeshthadeva is here generalising the semiring of based positive integers to the ring of polynomials over the ring of integers. More generally, the modern reader will also recognise what this text from the early 16th century is conveying as a pretty good characterisation of a polynomial.

Yuktibhāṣā illustrates the notation and its utility by means of the expressions which arise in connection with the first guess for the remainder $r_1(j)$. In computing the error $r_1(j-2)^{-1} + r_1(j)^{-1} - j^{-1}$, the common denominator is $(2j-2)(2j+2)j = 4j^3 - 4j = |4|0|\bar{4}|0|$ and the three terms in the error are displayed as

$$\frac{2j^2 + 2j}{4j^3 - 4j} = \frac{|2|2|0|}{|4|0|\bar{4}|0|}$$

and so on. The text actually contains these symbolic formulae in addition to a description in words, rather like, again, the Bakhshali manuscript (a thousand years on, there is still no confidence in symbolic methods?). Rules of algebraic operations can easily be formulated as rules for the entries in the boxes (as in the last sentence of the quotation above: for addition, just add the entries in the corresponding boxes). One interesting rule which *Yuktibhāṣā* mentions concerns cancellation of common factors: for example,

$$\frac{|a|b|0|}{|c|0|d|0|} = \frac{|a|b|}{|c|0|d|}.$$

It is not clear whether Jyeshthadeva meant the notation merely to bring out the parallel with decimal numbers or planned to exploit it as a full-fledged computational aid – aside from the astronomical Part II of *Yuktibhāṣā*, which uses very little of the pure mathematics of Chapters 6 and 7 of Part I, he only wrote this one book as far as we know. The terminology and the notation definitely had an influence on his own thinking about polynomials. For instance, while rejecting the unit correction to r_1 (see the last paragraph of the previous section), the reason he gives is not that the grossness is of order j^{-2} as I have stated there, but that it has now occupied the "place of the variable" (*rāśisthānam*) instead of the "place of units" (*rūpasthānam*, the constant term) as earlier. And it is futile to wonder what future mathematicians might have

made of these conceptual and notational innovations:[2] there was no room in Shankara's stylised verse writings for a compact symbolic notation and, after him, there was not much future left.

The other intriguing fact is that the analogy was not extended to include power series as was done by Newton, particularly in the light of Jyeshthadeva's declaration at the beginning of the book that there is no end to numbers. We can only speculate. Perhaps reservations about the unnamable still lingered; perhaps the effectiveness of abstract algebraic thinking in the handling of problems in geometry, familiar since the beginnings of mathematics, was not fully internalised; or, perhaps, the idea that a polynomial or a power series could be formally and abstractly *defined* for a general real variable such as an angle just as readily as for an integral variable, a circumstance that Newton made good use of, was not grasped.

13.3 Higher Order Corrections

From now on I revert to the standard notation. I will also express the error at any stage of the refining in terms of its behaviour for large j, rather than by all the places it occupies when computed exactly, as *Yuktibhāṣā* does. The reason is that while the exact computation of the error is easily done in the next stage (resulting, in our notation, in r_2), it is pronouncedly more tedious already in the next higher order of *saṃskāram*; asymptotic estimates have the same essential content for our present purpose and are somewhat easier to work out. In fact, while *Yuktibhāṣā* does cite the full formula for the next approximation r_3 in the form of a Sanskrit verse, it does not annotate it in Malayalam, gives no explanation for it (unlike for r_2) and does not provide any sort of supporting argument for its validity: there is neither *yukti* nor *bhāṣā* here. And it stops at that point; there is nothing in it about possible corrections of still higher order, no hint of whether they were even thought about. In trying to reconstruct these missing details it is of some help to be able to fall back on the notation (and the techniques) which the modern reader will be familiar with while not betraying the mathematical spirit of what is in *Yuktibhāṣā*, especially so in view of the fact that the final paragraphs of Chapter 6 dealing with the estimation of the remainder is written considerably less confidently than the rest of the book. Anything that will help render it less opaque should be welcome.

Yuktibhāṣā gives an elaborate explanation of how to refine $r_1(j)$ to $r_2(j)$ so as to reduce the error by two powers of j^{-1}, from j^{-3} to j^{-5}. The argument is somewhat convoluted in its logic but it works. For that reason, I first write down the suggested correction function and verify the validity of the claim before

[2] Perhaps it is not so futile. The preferred modern *definition* of a polynomial in one variable (as given in several standard textbooks on algebra) is essentially the same as Jyeshthadeva's: as an infinite sequence of coefficients, all but a finite number of which are 0, satisfying the rules appropriate for addition and multiplication. Modern *notation*, of course, does not follow Jyeshthadeva; it explicitly indicates the variable and its powers because such a functional notation generalises naturally to polynomials in several variables.

going into its motivation. It is convenient in describing this and subsequent *saṃskāram*(s) to denote the left side of the remainder identity by $R^{-1}(j)$:

$$\frac{1}{R(j)} = \frac{1}{r(j-2)} + \frac{1}{r(j)}$$

and by R_1^{-1}, R_2^{-1}, etc. the corresponding expressions when r is substituted by the successive approximations r_1, r_2, etc. Thus for $r = r_1$ we have

$$\frac{1}{R_1(j)} = \frac{1}{2(j-1)} + \frac{1}{2(j+1)} = \frac{1}{j - \frac{1}{j}} = \frac{1}{j} + \frac{1}{j^3} + \cdots .$$

The suggested second approximation is

$$\frac{1}{r_2(j)} = \frac{1}{2(j+1) + \frac{4}{2(j+1)}} = \frac{1}{2} \cdot \frac{1}{j+1+\frac{1}{j+1}}$$

which can be expanded as the geometric series

$$\frac{1}{r_2(j)} = \frac{1}{2}\left(\frac{1}{j+1} - \frac{1}{(j+1)^3} + \frac{1}{(j+1)^5} - \cdots\right).$$

The absence of even inverse powers in the expansion follows from the functional form of r_2, independently of the values of the coefficients of the two terms. The value of r_2 at $j-2$ is, correspondingly, given by

$$\frac{1}{r_2(j-2)} = \frac{1}{2} \cdot \frac{1}{j-1+\frac{1}{j-1}} = \frac{1}{2}\left(\frac{1}{j-1} - \frac{1}{(j-1)^3} + \cdots\right)$$

and we have the expansion

$$\frac{1}{R_2(j)} = \frac{1}{2}\left(\left(\frac{1}{j-1} + \frac{1}{j+1}\right) - \left(\frac{1}{(j-1)^3} + \frac{1}{(j+1)^3}\right) + \cdots\right).$$

The first term within brackets is the same as $R_1(j)^{-1} = j^{-1} + j^{-3} + \cdots$ of which the subleading term j^{-3} cancels the leading term in the expansion of the next bracketed term. So we have the desired result that $R_2(j)^{-1}$ differs from the exact $R(j)^{-1}$ by terms of order j^{-5} and smaller. The numerical coefficient 4 in the correction $4/(2j+2) = r_2(j) - r_1(j)$ is critical for the cancellation to work.

Yuktibhāṣā's explanation of how r_2 was arrived at is long but it is not very enlightening, effectively no better than guesswork and verification; the critical step is just stated: "To the earlier *saṃskāram* denominator $(r_1(j))$ must be added 4 divided by itself", followed by "so spoke the master (*ācāryan*)".[3] It then goes on to restate the prescription in the standard rational function form,

$$\frac{1}{r_2(j)} = \frac{\frac{1}{2}(j+1)}{(j+1)^2+1},$$

[3]In the context, the master can only be Madhava. This is the only place in the whole book in which Jyeshthadeva invokes authority in place of a reasoned justification.

as a verse in sanskrit, also almost certainly due to Madhava. The feeling is inescapable that Jyeshthadeva would have liked to provide a better *yukti* than he actually gives.

The fact that *Yuktibhāṣā*'s detailed description of the process results in a two-term continued fraction expression for $r_2(j)^{-1}$ turns out to be of great value when we attempt to reconstruct a *yukti* for it. And it becomes even more valuable for the historian, accompanied as it is by Madhava's verse expressing it alternatively as a simple rational function, in the light of a third, finer, approximation $r_3(j)^{-1}$ that *Yuktibhāṣā* proposes, again in a Sanskrit verse, again presumably attributable to Madhava.[4] This last correction is brought in very abruptly, without preamble or explanation: "Next [I] state a *saṃskāram* which is more accurate than the earlier one, to be applied after division by the odd number (j)". The sentence is followed by the formulaic verse for $r_3(j)^{-1}$, whose literal translation gives the formula

$$\frac{1}{r_3(j)} = \frac{((j+1)/2)^2 + 1}{\{4[((j+1)/2)^2 + 1] + 1\}(j+1)/2},$$

and the single word *iti*, indicating that the verse is a quotation. And, at this point, the whole discussion (and Chapter 6 itself) comes to a stop; no explanation of any sort is offered, no indication of where $r_3(j)^{-1}$ came from, no estimation of its accuracy. The last point however is a matter of direct but tedious verification: an expansion of $r_3(j)^{-1}$ in powers of j^{-1} leads to

$$\frac{1}{r_3(j)} = \frac{1}{j} + \frac{36}{j^7} + \cdots ,$$

vindicating the purpose for which the formula was devised.

In the absence of any textual guidance towards a *yukti* for the formula, we are on our own in trying to supply one, with no guarantee that any reconstructed proof will be faithful to what Madhava had in mind. But since it is the third in a sequence of approximations to the exact $r(j)^{-1}$, it is reasonable to expect that there was a proof for it which covered the second approximation as well (the first is easy, just inspection). The problem is that an extension of the rationale that gives the correct r_2 – cancellation of the subleading power of j^{-1} against the leading power in the correction – beomes quite involved. Perhaps that is how Madhava guessed the result. That will be disappointing as it is not supported by a good theoretical foundation; it would have been better if there was a proof that fits into the general philosophy of iterated refining, for example. Such a proof, faithful to the Nila style, has so far eluded commentators and we will return later to the question of where the difficulty might lie (and what might have discouraged Jyeshthadeva from attempting a proof at all).

[4]These anonymous verses occur unchanged in Shankara's self-styled commentary *Yuktidīpikā* on Nilakantha's *Tantrasaṃgraha*. Chronology rules out Shankara as their author. Tampuran-Ayyar thought that they occurred in *Tantrasaṃgraha* but they were almost certainly misled by interpolations in the manuscript they relied on.

What we do have is a modern proof that in fact establishes an infinite continued fraction expansion for the remainder (see below), presented in a very modern style, whose elements are not easily identifiable with a traditional *saṃskāram*. It builds on an elementary but brilliant observation of Tampuran and Ayyar: they turned the formula for r_3^{-1} given in the quoted verse, unpromising though it may be at first sight, into one with a strikingly regular structure. Because of its importance, I reproduce it here step by step. After a preliminary simplification, we have

$$\frac{1}{r_3(j)} = \frac{j^2+2j+5}{(2j+2)(j^2+2j+6)} = \frac{1}{(2j+2)\left(\frac{1+j^2+2j+5}{j^2+2j+5}\right)} = \frac{1}{(2j+2)\left(1+\frac{1}{j^2+2j+5}\right)}.$$

The denominator on the extreme right ($= r_3(j)$) is now manipulated as follows:

$$\begin{aligned} r_3(j) &= 2j+2+\frac{2}{\left(\frac{j^2+2j+5}{j+1}\right)} = 2j+2+\frac{2}{\left(\frac{(j+1)^2+4}{j+1}\right)} = 2j+2+\frac{2}{j+1+\frac{4}{j+1}} \\ &= 2j+2+\frac{4}{2j+2+\frac{16}{2j+2}}, \end{aligned}$$

resulting in the beautiful 3-step continued fraction expression for $r_3(j)^{-1}$

$$\frac{1}{r_3(j)} = \frac{1}{(2j+2)+}\,\frac{4}{(2j+2)+}\,\frac{16}{(2j+2)}$$

or, cancelling common factors (i.e., stopping at the last but one step),

$$\frac{1}{r_3(j)} = \frac{1}{2}\cdot\frac{1}{(j+1)+}\,\frac{1}{(j+1)+}\,\frac{4}{(j+1)}.$$

The first two approximants of this finite continued fraction are precisely *Yuktibhāṣā*'s first two approximations to $r(j)^{-1}$.

The fact that Tampuran and Ayyar work out the final correction as a formula with a structure generalising the first two approximations raises an interesting question. What made them look for a transformation of the unenlightening formula of the verse into a natural generalisation of $r_2(j)^{-1}$ in its continued fraction form: great insight or inside knowledge? – each step may be easy but they had to know in advance what they were aiming at. Given the historical continuity of astro-mathematical scholarship in Kerala, even if in an attenuated form in more modern times, it is not totally out of the question that the commentators had prior knowledge of the continued fraction presentation of $r_3(j)^{-1}$, perhaps transmitted along a parallel lineage not represented in a written text.[5]

[5]There is at least one precedent, a celebrated one, of a line of development bypassing an otherwise well-informed master: Bhaskara II not realising that Brahmagupta's *viṣama* quadrilaterals were restricted to be cyclic, see Chapter 8.4. It is pertinent to remember that

In any case, the commentary of Tampuran and Ayyar does have a short paragraph suggesting how the continued fraction formula for $r_3(j)^{-1}$ may have been derived by analogy with Jyeshthadeva's justification of $r_2(j)^{-1}$: start by computing the 'error'

$$\frac{1}{j} - \frac{1}{R_2(j)} = \frac{16}{4j^5 + 16j}$$

and adjust the correction to cancel the 5th power of j in the denominator; the details are extremely opaque, amounting to no more than guesswork followed by verification. Indeed, had Madhava or any of his followers wished to compute the next higher correction (let alone corrections *ad infinitum*) in the same way, they are likely to have been discouraged by the rapidly increasing algebraic complexity of the undertaking. It may be with good reason that no Nila mathematician cites corrections beyond the third.

The intriguing possibility remains: if Madhava did know the continued fraction form of r_3^{-1} and if he had Newton's instinct for "observing analogies" (see the quotation in Chapter 12.4), might he have conjectured its extension to general order as, for example,

$$\frac{1}{r(j)} = \frac{1^2}{2(j+1)+} \; \frac{2^2}{2(j+1)+} \; \frac{4^2}{2(j+1)+} \; \frac{6^2}{2(j+1)+} \cdots ?$$

It is a beguiling coincidence that the person who actually wrote down this particular expansion first was none other than the Newton scholar Whiteside. As stated above, it has a proof[6] and, not surprisingly, it is entirely unclear how this modern proof can be fitted into the mathematical culture of the Nila school. The question of how Madhava himself might have arrived at his formula for $r_3(j)^{-1}$ therefore remains open. More sharply: in the light of its recursive structure, can we reconstruct a *yukti* which is in tune with the many instances of the method of *saṃskāram* that we have met in the Nila corpus?

I describe now an attempt at a scheme of refining which is very natural from the *saṃskāram* point of view and produces $r_2(j)^{-1}$ from *Yuktibhāṣā*'s $r_1(j)^{-1} = 1/(2j+2)$. But its iteration fails in producing the required $r_3(j)^{-1}$ and the failure has a lesson about the limitations of 'trial and error' *saṃskāram*, a point to which we shall return after working out the details. The computation

---

Rama Varma Tampuran traced his descent from a line of royals that produced experts in the traditional sciences over several generations (and he was not the last one of the line). continued fractions were very familiar mathematical territory in Kerala, especially in relation to *kuṭṭaka*. Indeed, in a computation illustrating the efficiency of the third correction, Tampuran and Ayyar themselves present the intermediate numerical steps using the traditional tabular organisation of *kuṭṭaka*.

[6] According to C. T. Rajagopal and M. S. Rangachari in "On an Untapped Source of Medieval Keralese Mathematics", Arch. Hist. Exact Sci., vol. **18** (1978), p. 89, Whiteside communicated to them a proof, not reproduced in the paper and, as far as I know, unpublished under his own name (though he refers to it in a 2003 discussion on the Internet). Interested readers can consult a later paper of Rajagopal and Rangachari, "On Medieval Kerala Mathematics", Arch. Hist. Exact Sci., vol. **35** (1986), p. 91, which gives a proof in the modern style, not easy to follow, whose authorship is not explicitly attributed to Whiteside.

is straightforward. Start with $r_1(j) = 2(j+1)$ as the first guess and consider, as in most recursive approximation schemes, its difference from the exact $r(j)$ (the recourse to reciprocals is in view of the continued fraction expression that we are after):

$$\frac{1}{r(j)} = \frac{1}{2(j+1) + \frac{1}{x(j)}}$$

for some function x about which the only thing we know is the functional equation satisfied by r:

$$\frac{1}{j} = \frac{1}{R(j)} = \frac{1}{2(j-1) + \frac{1}{x(j-2)}} + \frac{1}{2(j+1) + \frac{1}{x(j)}}.$$

The two terms on the right add up to

$$\frac{1}{R(j)} = \frac{4j + \frac{1}{X(j)}}{4(j^2-1) + \frac{2j}{X(j)} + 2\left(\frac{1}{x(j-2)} - \frac{1}{x(j)}\right) + \frac{1}{x(j-2)x(j)}}$$

where X is defined, in analogy with R, by $X(j)^{-1} = x(j-2)^{-1} + x(j)^{-1}$. This equation does not, of course, determine X in terms of R. But there are obvious and natural approximations available, entirely in line with the *saṃskāram* philosophy: in the denominator, neglect the last two terms; since $x(j)^{-1}$ is a correction, they are putatively smaller than the first two terms and they both vanish asymptotically. (Such linearisations are often part of the approximations employed in *saṃskāram* algorithms, for example the Bakhshali square root, Chapter 6.5). We are left with a linear equation for $X(j)^{-1}$:

$$\frac{1}{X(j)}(2j - R(j)) = 4(jR(j) - j^2 + 1)$$

which, on substituting $R(j) = j$, becomes

$$\frac{1}{x(j-2)} + \frac{1}{x(j)} = \frac{4}{j}.$$

Thus, within the approximation, $(4x)^{-1}$ satisfies the same functional equation as r^{-1}, whose first-guess solution x_1 we already know:

$$\frac{1}{x_1(j)} = \frac{4}{r_1(j)} = \frac{4}{2(j+1)}.$$

Not only is the result of the first stage of refining,

$$\frac{1}{r_2(j)} = \frac{1}{2(j+1) + \frac{1}{x_1(j)}},$$

the same as in *Yuktibhāṣā*; the fact that the correction satisfies the same equation as the exact remainder but for the numerical coefficient, 4 instead of 1,

makes its iteration a trivial matter. To correct further for the grossness arising from replacing the exact x by x_1, write, in turn,

$$\frac{1}{x(j)} = \frac{4}{2(j+1) + \frac{1}{y(j)}}$$

and evaluate $Y(j)^{-1} = y(j-2)^{-1} + y(j)^{-1}$. In the 'linear approximation' for Y^{-1}, we have then

$$\frac{1}{Y(j)}(2j - 4X(j)) = 4(4jX(j) - j^2 + 1).$$

On substituting $X(j)$ by $X_1(j) = j/4$, it follows that y^{-1} satisfies the same functional equation as x^{-1}, with the same first-guess solution

$$\frac{1}{y_1(j)} = \frac{1}{x_1(j)} = \frac{4}{2(j+1)}.$$

This is where the problem lies: the *saṃskāram* has resulted in an expression for the third approximation,

$$\frac{1}{r_3'(j)} = \frac{1}{2(j+1)+}\;\frac{4}{2(j+1)+}\;\frac{4}{2(j+1)},$$

that deviates from the required $r_3(j)^{-1}$ in the numerator of the third 'term' of the continued fraction, 4 instead of 16; in particular, it will not compensate for the j^{-5} behaviour arising from the first two terms.

It is easy to see that the procedure can be repeated, resulting in the same correction, $4/2(j+1)$, at every stage, instead of the expansion constructed by Whiteside.

Disappointing as the outcome of this exercise has turned out to be, there is an important lesson here regarding the value of *saṃskāram*, when indefinitely extended, as a method of proof. As a general rule, no attention is given in such iterative computations to controlling the accumulated error resulting from the approximation involved in each step – it is a difficult problem – the geometric series and the sine/cosine series being exceptions. Consequently, there is no guarantee that higher order corrections will have the form required for them to lead to the correct convergent expansion when extended indefinitely. It is quite remarkable that the procedure Madhava adopts achieves the primary aim, that of ensuring that successive corrections lead to an improvement in the accuracy of $R(j)^{-1}$ by two powers of j^{-1}, even if it meant discarding a favourite technical device (at least in the simple version I have described above).

13.4 Variations on the π series

The basic π series is barely convergent: the sum of terms up to $j = 55$ (for this choice, see below) is off in the second decimal place. The corrections described

above overcome this handicap spectacularly efficiently, as we would expect from their design. We can get a rough idea of how good they are from the fact that the third correction is accurate to j^{-7}: $R_3(j)^{-1} = j^{-1} + O(j^{-7})$. For a j of about 10 (truncation after five terms), the error is of the order of one part in 10^7 and for Aryabhata's accuracy of four decimal places, R_3 applied after the first three terms will suffice. Tampuran and Ayyar in fact compute R_1, R_2 and R_3 (or rather their inverses diminished by j^{-1}, the errors) for $j = 55$ to get the circumference for a diameter of 10^{10} to 11 significant figures; the corresponding value of π resulting from the application of the third correction is accurate at that level, as we can see right away: 55^{-7} is roughly 10^{-12}.

The numbers for the diameter(10^{10}) and the corresponding circumference (31415926536 after the final correction) are both given by Tampuran and Ayyar in ingeniously constructed *kaṭapayādi* phrases which have every appearance of having been passed on from earlier times – it was not unusual, already in Madhava's time, to call at least the round numbers like 10^{10}, by their names, here *kharva* – though no reference is given; it would seem that the value $j = 55$ was not a random number picked just as an illustration. On the other hand, Madhava himself, according to Shankara's *Kriyākramakarī*, computed π to one more significant figure through the value of the circumference (expressed in *bhūtasaṃkhyā*) for a diameter of 9×10^{11} (expressed through its name, *nava nikharva*; why not the more convenient and natural 10^{12}?): $\pi = 3.14159265359$. It may well be that the enormous reduction in computation brought about by Madhava's approximation scheme led to several attempts at pinning down that elusive number, the exact value of π, the last such being Shankara Varman's (to 18 significant figures), as late as the 1820s. The calculations still involve a fair amount of labour; Tampuran and Ayyar for example had to compute every individual term, $1/j$ up to $j = 55$, to an accuracy of at least one part in 10^{11}.

No such numerical computation of π is given in *Yuktibhāṣā*. Moreover, extreme precision was not of any practical value to the astronomers, given the limitations of the means of astronomical observation, mainly the naked eye (if the astrolabe was known, there is no mention of it in their voluminous writings): Aryabhata's 1000 years old π together with Madhava's second (or higher if necessary) order interpolation of sines would have met their needs (recall, once again, that the accuracy of the sine table depends on the precision with which π is known). The conclusion is inescapable that the concern with π was driven by wholly theoretical considerations: how can the exact ratio of the circumference to the diameter be numerically defined or, in practice, determined to arbitrary accuracy?

What Madhava does next with these corrections strongly reinforces the impression of a theoretical bias. In a short subsection interposed between the formulae for r_2^{-1} and r_3^{-1} and titled "Variations on the method of determining the circumference", *Yuktibhāṣā* uses them to transform the basic π series into infinite series that converge more rapidly (and in fact are absolutely convergent). The key idea, as surprisingly elementary as it is pretty, is that the functional equation satisfied by the remainder, $R(j)^{-1} = j^{-1}$, allows the replacement of

j^{-1} in the basic π series by $R(j)^{-1}$ for any j. To implement it, note first that the basic π series can be rewritten formally as

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \cdots - \frac{1}{r(1)} + \left(\frac{1}{r(1)} + \frac{1}{r(3)}\right) - \left(\frac{1}{r(3)} + \frac{1}{r(5)}\right) + \cdots$$

since the terms involving the remainders cancel in pairs. The cancellation of course holds for any function of j whatsoever, not just for the remainders. Now rearrange the series as

$$\frac{\pi}{4} = 1 - \frac{1}{r(1)} + \left(\frac{1}{r(1)} + \frac{1}{r(3)} - \frac{1}{3}\right) - \left(\frac{1}{r(3)} + \frac{1}{r(5)} - \frac{1}{5}\right) + \cdots .$$

When $r(j)^{-1}$ are the exact but unknown remainders, each term in this rearranged π series vanishes except for the first two, by virtue of the functional equation (and the first two terms just define $r(1)^{-1}$ as $1 - \pi/4$). That is not very helpful; instead, *Yuktibhāṣā* replaces the general term $R(j)^{-1} - j^{-1} = r(j-2)^{-1} + r(j)^{-1} - j^{-1}$ by its value $e_i(j)$, the error resulting from any of the approximations $r_i(j), i = 1, 2, 3$ (and, in principle, so on) for $r(j)$, (and $r(1)$ itself by $r_i(1)$) from the previous section:

$$\frac{\pi}{4} = 1 - \frac{1}{r_i(1)} + e_i(3) - e_i(5) + \cdots$$

for each i.

Since the rearranged π series is formally exact independently of whether r is the exact remainder function, in fact for any function r, the correctness of this sequence of series in terms of the error functions e_i as representing $\pi/4$ depends, from a modern perspective, only on i) the legitimacy of the rearrangement itself and, if it is legitimate, ii) the convergence properties of the resulting series. The results of the previous section show that convergence is not an issue: $|e_i(j)| \sim j^{-(2i+1)}$. On the other hand, rearranging conditionally convergent series such as the basic π series is a delicate matter as we know from much later European work (Riemann's rearrangement theorem). Nilakantha and Jyeshthadeva both gave thought to questions of convergence, the former in the specific context of the geometric series (see Chapter 10.4) and the latter in connection with the domain of convergence of the arc tangent series (Chapter 11.5), but we can hardly expect them to have had an understanding of the subtleties of conditional convergence. Setting that aside, Jyeshthadeva's understanding of the logic of the rearrangement is clearly expressed (the passage comes after the explanation of the second correction, but holds for all the corrections):

> When an approximate circumference is subjected to this *saṃskāram* and the error (*sthaulyam*, $e_2(j)$ in the context) found, and it is positive, then the error from the next odd number ($e_2(j+2)$) will make it less (or, will be subtracted from it) and [hence, the result] more accurate. If the *saṃskāram* is carried out again and again (for higher

> and higher j), the result will become [still more] accurate. So, do this *saṃskaram* from the very beginning (from $j = 1$); the circumference will become [very] accurate. This is the justification (*upapatti*).

The essential point is that the corrections are to be applied to all the terms, "from the beginning".

It is straightforward to write down the series resulting from the various estimates for $r(j)$ in the rearranged π series. Their structure is described in a set of individual verses in Sanskrit, again most probably of Madhava's composition (they occur without change in Shankara's writings as well), the first of which corresponds to $r_2(j) = 2(j+1) + 1/2(j+1)$. The error for general j is

$$e_2(j) = \frac{1}{R_2(j)} - \frac{1}{j} = \frac{1}{j + \frac{4}{j^3}} - \frac{1}{j} = \frac{-4}{j^5 + 4j}.$$

The first two terms $1 - r_2(1)^{-1}$ in the transformed series add up to $4/5$ which, as it happens, is equal to $4/(1^5 + 4 \cdot 1)$; we have thus

$$\frac{\pi}{16} = \frac{1}{1^5 + 4 \cdot 1} - \frac{1}{3^5 + 4 \cdot 3} + \frac{1}{5^5 + 4 \cdot 5} + \cdots .$$

The rearranged series corresponding to the first correction $r_1(j) = 2(j+1)$ is similarly obtained from the error function

$$e_1(j) = \frac{1}{R_1(j)} - \frac{1}{j} = \frac{1}{j^2 - 1} - \frac{1}{j} = \frac{1}{j^3 - j}:$$

$$\frac{\pi}{4} = \frac{3}{4} + \frac{1}{3^3 - 3} - \frac{1}{5^3 - 5} + \cdots .$$

The alternation in the sign of the term corresponding to a given j between the two series comes from the first and second corrections being of opposite signs; the resulting compensatory effect was noted in the prefatory explanation in *Yuktibhāṣā* of the general philosophy of repeated corrections.

The series corresponding to the third correction is not given in any text, probably because the numerical coefficients occurring in $e_3(j)$ are not so small;[7] these series probably were not meant for any practical use in any case. And, as though to illustrate the freedom in the choice of $r(j)$ in the rearranged series, the series corresponding to the '0th guess', the very first, rejected, trial value $r_0(j) = 2j$, is described in a verse. As is easily determined, the series is

$$\frac{\pi}{4} = \frac{1}{2} + \frac{1}{2^2 - 1} - \frac{1}{4^2 - 1} + \frac{1}{6^2 - 1} + \cdots .$$

In the background of the several variations of the π series the Nila mathematicians played around with, this particular version has no great intrinsic interest. But one can cut it off and estimate the remainder in much the same way as

[7] $e_3(j) = 36/(j^7 + 7j^5 + 28j^3 - 36)$.

for the basic π series and, as pointed out by Rajagopal and Rangachari (cited in the previous section), derive a continued fraction expansion found later by Euler.

For good measure, the last cited verse on the topic has no use for the error functions at all but simply regroups two adjacent terms into one; this most elementary form of rearrangement,

$$\frac{1}{j-2} - \frac{1}{j} = \frac{2}{j^2 - 2j} = \frac{2}{(j-1)^2 - 1}, \ j = 3, 5, \cdots,$$

leads to the two series

$$\frac{\pi}{4} = \left(1 - \frac{1}{3}\right) + \left(\frac{1}{5} - \frac{1}{7}\right) + \cdots = \frac{2}{2^2 - 1} + \frac{2}{6^2 - 1} + \cdots$$

and

$$\frac{\pi}{4} = 1 - \left(\frac{1}{3} - \frac{1}{5}\right) - \left(\frac{1}{7} - \frac{1}{9}\right) - \cdots = 1 - \frac{2}{4^2 - 1} - \frac{2}{8^2 - 1} - \cdots.$$

It is noted that all terms are positive in the first series and all but the first negative in the second series; they are not alternating series.

The material described in this section brings Chapter 6 of *Yuktibhāṣā* to a close (but for the single verse with the formula for the third correction). As in many other innovations in the Nila work, the basic idea of rearranging and reorganising recalls techniques from the past, in this case the very remote past, having little to do with mathematics: if these manipulations of an infinite series, ostensibly as a means of fixing the value of π, recalls to the reader's mind the regroupings, repetitions, substitutions, insertions etc. of the syllables of long Vedic texts – *padapāṭha* and *prātiśākhya*, see Chapter 1.3 – it is not entirely without reason. The recitative techniques had the ostensible aim of helping memorisation but it is difficult to read them – or, even better, to listen to them – without being struck by their infinite variability, their playfulness almost. Despite first impressions, the many rearrangements of the π series were not only, or even primarily, meant to produce sharper and sharper values of π; as we have seen, there were other, more efficient, means of doing so. It is better to view them, at least in part, as a shift away from usable mathematics that Chapter 6 of *Yuktibhāṣā* signals. The message from the most advanced and abstract parts of its content, from calculus to abstract algebra, is that the mathematics is no longer motivated only by its applications but more and more by its own inherent potentialities. The same message comes through in Chapter 7; the properties of the sine as a function on the circle, especially its representability as a power series, (or of the cyclic quadrilateral for that matter) have no direct use in astronomy.

Finally, a question which goes unmentioned in *Yuktibhāṣā*, that of the irrationality of π, particularly surprising in a chapter whose exclusive concern – and from several different angles – is the ratio of the circumference and the

diameter. Nilakantha's unambiguous assertion that they are incommensurable (see the quotation in Chapter 9.2) occurs in the *Āryabhaṭīyabhāṣya*, written at about the same time as *Yuktibhāṣā*. Nilakantha was not given to unsupported mathematical *obiter dicta*, especially not in the *Āryabhaṭīyabhāṣya*. What reasons might he have had for the conjecture?

To set the stage for a possible answer, let us briefly recapitulate a standard European proof, not the original one of Lambert but an application of a general criterion of irrationality formulated somewhat later (end of the 18th century) by Legendre. Suppose we have a positive numerical quantity u expressed as a continued fraction,

$$u = \frac{m_1}{n_1 + \frac{m_2}{n_2 + \cdots}},$$

m_i, n_i positive integers with $m_i \leq n_i$; so $u < 1$. Defining $v(> 0)$ by $u = m_1/(n_1 + v)$, we have $v = m_1/u - n_1$. Suppose now that u is rational, say a/b; then $a < b$ and $v = bm_1/a - n_1$ is (positive) rational, with denominator (=a) less than the denominator of u (=b). We can continue the steps by writing $v = m_2/(n_2 + w)$ and conclude that w is (positive) rational with denominator less than that of v; and so on. From the boundedness of positive integers below, it now follows that the continued fraction expansion must terminate. Contrarily, if u has an infinite continued fraction expansion meeting the conditions stated above, then u is irrational.

Whiteside's extension of the first three corrections to an infinite continued fraction expansion of $\pi/4$ meets the sufficient conditions (they can in fact be weakened) and hence proves its irrationality.

Could Nilakantha (or someone before him, say Madhava) have constructed such a *yukti* or one similar to it? Of the two components that are vital to the proof, one mathematical and the other metamathematical, the former, namely the use of continued fractions, was very much a part of the Nila ethos. Apart from their occurrence in the truncation remainders in the π series itself, the method of descent that is employed (and the accompanying reduction to smaller and smaller denominators) is classically Indian, reminiscent of *kuṭṭaka*.

The metamathematical component is of course the reliance on *reductio ad absurdum* and that is a more problematic issue. There are no *reductio* proofs in the mathematics of India – excluding one late and feeble attempt to show that negative numbers cannot have square roots since all squares are positive – though the status of the principle of the excluded middle as a logical tool was endlessly debated by philosophers. Unlike in the hairsplittings indulged in by them, generally encompassing propositions far removed from the precision of mathematical statements, the irrationality issue is sharply framed: if the supposed rationality of π implies the finiteness of its continued fraction expansion(s) (meeting certain conditions), can it be concluded that the fact that the expansion (again, meeting certain conditions) is infinite implies that π is irrational? Nilakantha certainly would not have had a problem establishing the direct implication: rationality $\rightarrow$ finite expansion. As regards the reverse implication, given the philosophical acuity and heterodoxy of his last writings

– among them the *Āryabhaṭīyabhāṣya* – he may well have been convinced of its correctness, even if he kept his reasoning to himself. There are other examples in the Nila corpus of disregard for philosophical dictates, for example Jyeshthadeva's acceptance of unboundedly and unnamably large powers of 10 in the approach to the infinitesimal, after having said that there is no end to numbers *because* there is no end to their names: effectively, that to name is to bring into being.

It needs to be clearly said, again, that these remarks on the irrationality question are not based on any documentary evidence; they are at best a not wildly implausible attempt to get inside Nilakantha's mind, at worst pure unsupported speculation. Perhaps his conjecture was just an insight born out of frustration at being unable to find a finite numerical expression for π.

Within a hundred years of the writing of *Yuktibhāṣā*, creative mathematical activity came to an end, in the Nila basin, in Kerala and in India as a whole. When Srinivasa Ramanujan emerged on the scene three centuries later, and on the opposite coast no more than a few hundred kilometers from the Nila, the mathematics he was brought up on was that of the European tradition, transformed almost unrecognisably within the same three centuries into *the* mathematical culture of the world. Nothing we know of Ramanujan tells us that he knew of the still older legacy that was his.

Part IV

Connections

14

What is Indian about the Mathematics of India?

14.1 The Geography of Indian Mathematics

In contrast to the first three Parts of this book, this concluding Part IV has little that is strictly mathematical and nothing that is new; it is, instead, a synthesis of the more prominent of the many facets of the mathematics of India that give it a distinctive cultural personality of its own. We have already come across quite a few of them in their isolated settings but our primary focus in the coverage until now has been on the mathematics itself (and, to a modest extent, its historical background). And we have seen that that is wholly in tune with the widely accepted view that at its core the discipline of mathematics – the finest attestation of the capacity of the human mind to reason things through "minutely and precisely" (to invoke the gift attributed to Agni the divine enumerator in the *Ṛgveda*) – is largely free of the cultural context. The total mathematical profile of a civilisation, however, encompasses more than the universal and immutable mathematical truths it discovered for itself, more than can be gathered together in a neat collection of theorems. Questions concerning such intangibles as the value attached to mathematical activity, the manner in which new directions of mathematical enquiry were picked (or picked themselves), how mathematical ideas were motivated, pursued and made use of, in other words the philosophical and cultural whole of which mathematics was a part, seem to have not always the same answers in different cultures.

Of the mathematical cultures of the ancient world, the one that has had the greatest attention given to it, leading to a corresponding depth of understanding, is Hellenic mathematics, and for excellent reasons. Ancient Greece and the contiguous Mediterranean regions produced not only a wealth of mathematics but also finely detailed mathematical texts (and the philosophical

P. P. Divakaran, *The Mathematics of India*, Sources and Studies in the History of Mathematics and Physical Sciences, https://doi.org/10.1007/978-981-13-1774-3_14

writings that underpinned it all) most of which survived into modern times. The birth of modern mathematics in Europe in 16th-17th centuries was directly inspired by and modelled on its Hellenic precedents, and it has grown rapidly into the extraordinarily fruitful activity that is today's mainstream mathematics. But hardly any change has been found necessary in its original epistemic foundations or in the general intellectual matrix in which it is embedded; no serious philosophical recalibration has been required (at least until recently) to fit the latest advances into the mathematical world view of the Greeks or, conversely, to grasp what the Greeks achieved from today's perspective. The situation in regard to the reading and interpretation of traditional Indian mathematics could not be more different. Modern mathematicians (including those from India) and historians, brought up as they are on Euclid and Archimedes and the European inheritors of their legacy, and wishing to locate the mathematics of India in *its* natural context, have first to recreate that context. The mathematics itself is not a mystery; but if scholars are occasionally sceptical of, say, the logical credentials of the discoveries of a Baudhayana or an Aryabhata or a Madhava – did they really *prove* the results ascribed to them? – it is a perfectly understandable response. More generally, the linkages of Indian mathematical life to the total culture that nourished it, such as faith and ritual, philosophical predispositions and, especially, language (mainly Sanskrit) in its many manifestations, were much more intimate than they are today. Especially over the past few centuries, our capacity to reimagine the intellectual world in which Indian mathematics flourished has been greatly attenuated as our connection to that world became more and more remote. At a time when Europe was discovering its intellectual roots, India was beginning to lose contact with its; it is only very recently, with the weakening of the colonially induced disregard of its own traditions of learning and teaching, that a start has been made in the critical – and, sometimes, not very critical – reevaluation of its mathematical past.

There is a straightforward but still useful sense in which the term "the mathematics of India" – the title of this book – can be understood: the totality of mathematical activity of the inhabitants of cultural India over its recorded history. From the first formulation of orally expressed decimal enumeration, before the compilation of the *Ṛgveda*, to its decline and disappearance in the 17th century, the pursuit of mathematics remained an unbroken and constantly renewed part of India's culture: if we exclude the largely undeciphered mathematical messages from the Harappan civilisation from consideration, that still makes the story more than three millennia old. There really were no significant gaps in this long history. Even during the poorly documented – as far as dominantly mathematical texts are concerned – interregnum spanning the half millennium between the earliest recensions of the *Śulbasūtra* and the combinatorics of Pingala and Bharata, mathematical life was far from extinct: new versions of the *Śulbasūtra* continued to be composed, the later Brahmanas and the Upanishads did not stop referring to numbers and their manipulation, Jaina and Buddhist cosmogonists took up the exploration of very large numbers where

the *Taittirīyasaṃhitā* left off, and so on. No other civilisation except, maybe, that of China can claim a comparably vigorous mathematical longevity. And if we push the history back another thousand years to the Indus Valley period, it is an exciting prospect – of which the first signs are beginning to come to light in geometry (see Chapter 3.3) – that the dark ages between the end of its urban phase and the beginning of the Vedic period proper may soon reveal some of their secrets.

The historical continuity goes hand in hand with an equally remarkable but less immediately obvious geographical continuity. At some time or the other in this long history, the culture of mathematics has touched virtually every part of the geography of India. Since the only way knowledge could be carried across the vast expanses of the country was in the minds of those who travelled, the story of the diffusion of mathematical knowledge is also the story of the movement of people, voluntary or forced. We have seen examples of these migrations in the earlier parts of this book but it is dramatic to look at the main currents at one go. To begin with, there is the resurfacing of the Indus Valley counting and measurement systems in deep south India at an unknown time, probably well before the beginning of the common era. During the same first millennium BCE, we have the spread of the orally literate Vedic culture eastwards from modern Panjab and the upper reaches of the river Yamuna (where the earliest *Śulbasūtra* were composed; 1000 BCE or soon after) to modern Bihar (where the final form of the Sanskrit syllabary came to be fixed and the analytic *padapāṭha* techniques were perfected, thanks to which we can say with certainty that the mathematical information we can extract from the Vedic texts is authentic; around 500 BCE). These are long journeys from long ago, but the circumstantial evidence for them is quite persuasive and, as regards the southern migration from the Indus valley, getting to be more so.

Coming to more historical and better documented times, the next defining event was the rise of Taxila as the crossroads of several cultures, Indian and alien, all merging into a composite whole, and as one of the major centres of learning of the ancient world. Taxila's preeminence as a place of scholarship, and of the Gandharan region as a whole, got a big boost after the invasion of Alexander when it began to open up to influences from the Hellenic world (Chapter 6.1) culminating, as far as mathematics is concerned, in the adoption of the idea of astronomy as a mathematical science. Greek astronomy came to India in a mature form, probably after Ptolemy. Within another century or so, when the Siddhantas – wholly Indian geographically and culturally for all the debt they owed to the Greeks – began to be written, the locus of astronomical activity shifted decisively to Avanti and Malava, 1000 km and more to the southeast. There is little room for doubting that this particular displacement was driven by the fact of Ujjain lying very near the Tropic of Cancer; the longitude through it was picked as the prime meridian – though there is no good scientific reason for doing so – a distinction it retained as long as Indian astronomy remained a living science and acknowledged as far south as the 9 degrees latitude of Kollam.

The Siddhanta phase came to a close with Varahamihira.[1] But before that, Aryabhata had announced himself in Kusumapura, somewhere near Nalanda, precise location unknown, far to the east of Ujjain. As I have argued at length in Chapter 6.4, he or his immediate ancestors were very likely refugees forced to flee across almost the full breadth of the north Indian plains, part of a shift of the intellectual centre of gravity from northwest to eastern India where numerous centres of learning came up, Nalanda being the most famous of them. What is astonishing is that, within three or four decades of its writing, Varahamihira in Ujjain came to know Aryabhata's work in such detail as to be able to integrate his trigonometry in what I have called the independent chapters of *Pañcasiddhāntikā* (see Chapter 6.2) and to critcise him by name for proposing the midnight doctrine; Aryabhata, almost certainly, was still living.

From Aryabhata onwards, the chronological sequence of mathematical events is, broadly speaking, unbroken though there are many not-so-minor figures – Aryabhata II and Jayadeva from among those we have come across, for example – whose dates are known only conjecturally. Patterns of geographical continuity are less well established. Confining ourselves to just the big names, there is the first Bhaskara in Valabhi (modern north Gujarat) in the first half of the 7th century and his contemporary Brahmagupta not too far away, around Ujjain (probably). And then, apparently out of nowhere, Govindasvami and Shankaranarayana make their entry on the Kerala coast. They mark the very first inroads of Aryabhatan science, through the agency of Bhaskara I's writings, south of the Vindhya range, the first seeding of the ideas that later grew into the majestic corpus of the Nila school. We are largely ignorant of those who kept the tradition alive, but they were almost certainly part of, or descended from, the first settlements of Brahmins brought down the coast by local rulers (Chapters 8.1 and 9.3). About the same time, Mahavira is in residence in the capital of the Rashtrakuta kingdom in Karnataka; he certainly was no Aryabhatan and we know little about his immediate roots, mathematical or otherwise.

Later, Bhaskara II's ancestor made his much shorter trek across the Vindhya, from Malava to his natal village near Chalisgaon, a displacement forced by the raids of Mahmud of Ghazni, who scattered Hindus "as atoms of dust in all directions" (in Alberuni's words): "Hindu sciences have retired far away from those parts of the country conquered by us". Within a few decades of the death of Bhaskara II, whatever little was left of the mathematical life, in whichever part of north India, was extinguished; men of learning were not merely scattered as dust by the armies of Muhammad Ghori, many were put to the sword and the study centres razed and burned. The scholarly centre of gravity shifted once more, to south India again; the first ever commentary on the works of

[1]The name is an intriguing mixture of orthodox Sanskrit/Hindu (Varāha is the boar incarnation of Viṣṇu) and Persian/Zoroastrian (Mihira is the god of light or the Sun which, at that time, was also associated with the transcendental Mahayana form of the Buddha, Amitabha).

Bhaskara II (specifically on *Līlāvatī*) was probably Parameshvara's, written a thousand kilometers south of Chalisgaon.

The revival of astronomy and mathematics in Kerala inspired by Madhava cannot however be attributed solely to the chain of political-military events that convulsed the north Indian plains following the Ghori visitation. Madhava himself, of coastal Karnataka origin, may well have benefited from the support the Vijayanagar kingdom gave to Hindu scholarship, as did probably Narayana a generation or two earlier. But the community of Namputiris (and, a little later, their non-Namputiri offspring) that carried forward Madhava's legacy on the banks of the Nila was descended from the several waves of Brahmin migrations that began well before the 12th century. Whatever the reasons that started them off in the first place, their journeys were very long, from as far away as the upper Ganga basin according to their own chronicles, in many stages and over correspondingly long periods of time.

And when mathematical activity did revive in north India, it was brought about by a reverse migration, that of the Daivajña clan from Maharashtra to Varanasi (Benares), carrying with them the teachings of Bhaskara II.

This briefest of brief summaries is meant primarily to bring to the forefront of our minds the most significant stages in the evolution of Indian mathematics as a function of time and space. But just running through the names of the main protagonists and their theatres of action brings into sharp relief not just their personal contributions to that story but also the unity of the conceptual framework which held it all together and of the technical apparatus employed, over physical distances of thousands of kilometers and time scales of millennia. We have seen any number of instances of how the earliest insights were never lost sight of but became a vital part of new knowledge, often much later and at the other end of what is now called the subcontinent. Thus, the origins of Aryabhata's trigonometry and Brahmagupta's geometry of cyclic quadrilaterals, though they seem to have little in common on first acquaintance, are both to be found in the *Śulbasūtra*: the former in the orthogonality of the line of centres and the common chord of intersecting circles (which may very well have been inherited from Indus Valley geometry) together with the diagonal theorem and the latter in the fact that the rectangle of the diagonal theorem is the simplest pre-Brahmagupta (cyclic) quadrilateral. Subsequently, both these themes underwent further major transformations, the trigonometry in particular leading on to the infinitesimal calculus of the circle (with help from yet another ancient insight, that numbers have no end). Other examples are perhaps not so spectacular but remarkable nevertheless: the binomial coefficients that Pingala worked out for the classification of metres remained more or less dormant, mathematically speaking, for a very long time before reappearing in Narayana's general combinatorics and, soon after, in the coefficients of the sine series.

Part of the background to the pan-Indian sweep of the mathematical – more generally, intellectual – culture is the predilection Indians always had for long journeys, not always in flight from invading armies. Pilgrims and

proselytisers, teachers and those seeking learning, architects, sculptors and painters looking for patronage and employment, lots of people were on the move, and for diverse reasons. They recognised no political boundaries and those who had special skills to sell paid little attention to who ruled a particular patch of the country as long as it was receptive to what they had to offer.[2] Brahmins were particularly enthusiastic travellers, going wherever the supposed power of their rituals found a market. But they were also the custodians of traditional knowledge – especially after Buddhism disappeared from most regions of India – which went with them wherever they did.

The picture we have then is of an intellectual network covering most parts of India by about 1000 CE, with a strong sense of purpose and holding on zealously to its knowledge system and the language, Sanskrit, in which it was expressed and preserved. Every student, wherever he may have been living, began with the study of Sanskrit in all its aspects before those who had higher aspirations went on to more specialised areas of scholarship. Virtually all books were composed in a standard, functionally efficient, Sanskrit, in verse or prose; even *Yuktibhāṣā* in this sense is an almost-Sanskrit text, except that the narrative happens to be in an already heavily Sanskritised Malayalam. And every student was required to master in the course of his training the work of the great masters of the past, the *pūrvācārya*, of all times and from everywhere. Indian scholars did not have to rediscover the attainments of their predecessors as Europe had to rediscover those of Classical Greece before laying claim to its inheritance. It is not in the least surprising in this light that the majority of the manuscripts of the great mathematical classics were found, and are still being found from time to time, in Kerala: in addition to the increased popularity of writing (on palm leaf) as a means of text preservation by the 15th century, the decisive factor was that Kerala was the last sanctuary of a tradition of learning, the last node of a network that touched all of India and its entire history.

In a very real and physical sense, the story of mathematics in India is a story of journeys.

I end this bird's eye view of the Indian mathematical landscape with a brief look at the possible role the faith of individual mathematicians may have played in the way they approached their vocation. In brief, it was negligible. The geometry of the *Śulbasūtra* was of course dictated by (Hindu) Vedic ritual science and its architectural demands. (Decimal enumeration had to do with more mundane needs perhaps and, theoretically, with grammar rather than with ritual). That did not stop it from being reborn in a wholly secular setting (Aryabhata) and becoming the common heritage of Buddhist and Jaina mathematicians as well (who were also pioneers in exploring the infinitude of numbers, it may be recalled). Though the practice of mathematics became progressively a monopoly of Hindus – naturally so in view, especially, of the weakening of the Buddhist faith in India – there was also the odd non-Hindu

[2] It is amazing to realise that Kerala was never part of a kingdom or empire or any sort of political entity based outside its natural boundaries, not when Brahmins travelled there from farthest north India and not, as a whole, until it became part of independent India.

contributor to the making of the whole fabric, even a Muslim one if we grant guest status to Alberuni in recognition of the long years he spent in India. The Hindus themselves came in all shades of belief and non-belief: from the very orthodox Brahmagupta (his geometry was non-sectarian but he stuck obstinately to a post-Vedic value of π and believed that eclipses were caused by the sun and the moon being swallowed by demons) to the more open-minded Bhaskara II who, at the same time, was a devotee of a whole pantheon of gods – going by the benedictory verses – to Jyeshthadeva who bowed to no god at all. It is easy to accept that Nilakantha was a believer only by the accident of his birth, turning, in the course of a rewarding life, to a form of empirical rationalism and renouncing unthinking loyalty to *śruti* and *smṛti*. As for the two towering figures who anchored the historical phase of Indian mathematics at either end, we have seen in some detail that if Aryabhata was a Hindu, he was one only notionally and the evidence for that comes from his own words: he seeks no help from gods and goddesses to navigate the ocean of truth and falsehood, only that of the boat of his own intelligence, *svamatināvā*. And about Madhava we do not have enough information, in his words or in his followers', to go beyond a guess: if he was a follower of the *lokāyata* philosophy as the later members of the Kudallur (Sangamagrama) line were, he would not have been considered a good Hindu, let alone a good Brahmin.

The remarkable unity of thought and practice seen in the mathematics created by such a diverse collection of people – belonging to different places at different times and following different faiths or no faith at all – was quite exceptional for India. It was not the case in other areas of Indian intellectual life such as its philosophies and logics (the plural is deliberate) where finely drawn distinctions and passionate disagreements about them were the norm. Perhaps we have to thank the universality and absolutism of mathematical truth for it or, perhaps, the fact that mathematicians and astronomers seem to have learned to keep their distance from their argumentative philosopher brethren.

14.2 The Weight of the Oral Tradition

Of all the ways in which the cultural context determines the approach of a civilisation to the enterprise of doing and communicating mathematics, none is more singularly Indian than the absence of a written symbolic notation. As discussed at the beginning of this book (Chapter 1.3), this was a reflection of the more general repudiation of writing as a vehicle for the composition and transmission of all sacral texts, not only the Veda proper but also the extensive exegetic literature, including the *Kalpasūtra* of which the *Śulbasūtra* were a component. To modern mathematically literate readers, submerged in a proliferation of notational innovations as they are, it is not natural that a construction such as that of a square equal in area to a circle can be made intelligible without writing equations or even that the place-value principle can be implemented through words alone. Consequently, doubt is sometimes expressed

about whether writing was really unknown to or rejected by the Vedic seers – a position occasionally misattributed to the influence of Western linguists and philologists – or whether it was in use but on perishable materials.[3] The answer to the doubts, correspondingly, comes at two levels. The first is that mathematics can be done very effectively without symbols and formal equations (and transmitted orally) and the proof is simply that Indian mathematics managed quite satisfactorily without these artefacts of communication until its end, from the operations of basic arithmetic (see Chapter 5.3 for the example of long multiplication) to the sophisticated and intricate reasoning behind the sine series. For the present day scholar, reading narrative mathematics can be something of a challenge at first but every serious historian has had to turn verse and prose in Sanskrit or Malayalam into equations and other kinds of relations among symbols; it is an acquired skill, like learning a foreign language.

The less superficial response pertains not just to mathematics but to Vedic Sanskrit in all its aspects: the accumulation of numerous pieces of circumstantial (but strong) evidence described in Chapter 3.1, ranging from the organisation of the consonant grid on the basis of the appropriate sound-producing part of the whole vocalising apparatus (Chapter 0.4), not randomly as for instance in English, to the techniques employed in aid of memorisation and authentication. Especially relevant are specifically vocal and aural linguistic elements like the duration and pitch of a syllable and such sundry facts as the etymology of words like *adhyayana* (= the act of reciting → teaching) or *adhyāya* (= that which is recited → chapter, as in Panini's *Aṣṭa-adhyāyī*). It took a Buddhist work, finally, to acknowledge publicly that language could also be written down: *Lalitavistāra* with its reference to many scripts, whose core may be datable to the time of the first appearance of Brahmi writing.

Even then, the hold of the oral tradition was so strong in mathematical matters that when numbers occur in inscriptions in the very same Brahmi script, they are written as a faithful symbolic rendering of their names, with single symbols for powers of 10 and suffixes doing duty for grammatical rules of composition (Chapter 5.3). It is as though Romans were to have written C^{III} for 300: symbolic but not positional. Not coincidentally (I think), the appearance of writing is also synchronous, give or take a century, with the first mention (by Pingala) of zero as a number, not as a grammatical convenience as in Panini's *lopa*. The reason for the long delay between the fully evolved decimal number system of the *Ṛgveda* and its extension to include *śūnya* should be clear from our earlier discussion: nominal counting and basic arithmetic could very well do without a zero. The crucial fact is that the name of a number, unlike its positionally written collection of symbols, is a faithful realisation in words of

[3] There are other factors contributing to the scepticism, like disbelief in the ability of an average (or not so average, perhaps) human to memorise massive quantities of text or the general identification by historians of the advent of writing as an unfailing sign of civilisational progress. The opinion is sometimes expressed that the Vedic civilisation, being the most advanced not only of its own time, could not possibly have been ignorant of writing but, fortunately, it is not widely shared.

its polynomial representation; in both, a power of 10 whose coefficient is zero is simply dropped, *lopa* in action.

The impact of orality on the complementary notion of infinity was of a very different kind. We know from the *Taittirīyasaṃhitā* that the potential infinitude of numbers was an object of fascination from the earliest Vedic times (Chapter 5.3). In an oral culture, numbers had to have names for them to have a palpable identity and it is possible to argue that the problem of assigning names to higher and higher powers of 10 uniquely and unambiguously – an ideal which was never achieved in the numerous power lists – was what came in the way of an open acknowledgement of the endlessness of numbers.[4] Various stratagems were tried to overcome this epistemic hurdle – for it was that – from *Taittirīyasaṃhitā*'s *sarvasmai svāhā* as a bridge from the finite to the infinite (rather like the $\cdots$ of modern notation) to Prince Siddhartha's recursive trick in the *Lalitavistāra* (the number of grains of sand in as many riverbeds as there are grains of sand in one riverbed, see Chapter 5.3). The problem would have been particularly acute for followers of the linguistic philosophy of Bhartrhari (like the Kudallur Namputiris of the Nila country, the putative descendents of Madhava, see Chapter 9.4) for whom, loosely speaking, an abstract object became real through its name being uttered (ontological nominalism, if we must have a name for it). On the other hand, the inexhaustibility of lingustic expressions had been recognised earlier by Patanjali who relates a fable in which the preceptor of the gods, Bṛhaspati, tries to teach one of them, Indra, grammatical expressions by enumeration and gets nowhere in a thousand years of the gods. Patanjali also has the solution (he was no nominalist): " Some work containing general and particular rules must be composed"; Panini's *Aṣṭādhyāyī* in other words.

It is Panini's astonishingly deep insight, the deepest of many, that the ideal grammar should seek to regulate language by means of rules which are categorical. The number of categories at each hierarchical level in the organisation of the totality of linguistic expressions is finite and the rules that apply to and, therefore, define each category are the same for all members of the category (with exceptions made for a few Vedic usages; he was after all dealing with a living language which had evolved organically from the Vedic). But each category may, in general will, contain an unlimited number of members and it is enough to say that the appropriate rules will apply to all of them without having to name them individually. Thus, though the *sandhi* rules for vowels, $a + i \to e$ and $a + u \to o$, cover a potentially infinite number of examples, it is unnecessary to list them one by one like Brhaspati. It is equally astonishing that such a powerful epistemic principle, even better suited for mathematical reasoning than for grammatical, seems to have had an impact on Indian mathematicians only sporadically. Of the infrequent examples that come to mind, two are noteworthy: Brahmagupta's enunciation of the rule of multiplication of

[4]The 5th century AD Mahayana philosopher Vasubandhu, a contemporary of Aryabhata, speaks in one of his works of 60 named powers of 10 but actually lists only 52 of them, explaining that 8 (which 8?) have become *vismṛtam*, lost to memory.

positives and negatives covers all (real) numbers (as well as algebraic symbols); since symbols do not have a fixed intrinsic sign, it can only be understood as a metamathematical rule for operations on sets labelled, *à la* Panini, *dhana* and *ṛṇa* and that is how it is stated: "*ṛṇa* multiplied by *dhana* is *ṛṇa*, by *ṛṇa* is *dhana*, *dhana* multiplied by *dhana* is *dhana*". Much later, it became a key input in the categorical treatment of polynomials of general degree (with arbitrary integer coefficients) in *Yuktibhāṣā*, acknowledged by the citation of the relevant verse. Similarly, his formulation of the quadratic Diophantine problem is consciously general and applies to the whole (infinite) class of such equations for every choice of coefficients from the set of integers. And it is to qualify the number of solutions of one such equation that the word 'infinite' (*ananta*, unending) was used for the first time in a mathematical context.

In a philosophical sense, the irreconcilability of nominalism with the mathematician's need to deal with very large numbers was never resolved satisfactorily. But it did find a practical resolution in the course of the development of calculus by Madhava and the Nila school. We can recognise today that the central idea of the Nila solution to the problem of infinity – once again, our guide is *Yuktibhāṣā*, with help from Nilakantha's *Āryabhaṭīyabhāṣya* – lies in the notion, in modern terminology, of the limit of a sequence. Infinities are everywhere in the Nila calculus: in increasing order of sophistication, first as the number of terms in the various series, then as the number of steps in an inductive proof and, most critically, as the number of segments into which a geometric object such as an arc is divided in the approach to the infinitesimal limit (division by infinity). But Jyeshthadeva is a nominalist at heart, as proved by his dictum that we cannot *know* all the numbers *because* there is no end to their names. How then can he justify these infinite processes involving objects that are unknowable in principle since they cannot all be assigned names?

The honest answer is that he cannot, not as long as he remains faithful to his dictum. For the construction of the infinite series for π, he falls back on that old favourite, recursion, and computes the general term from its predecessor; though the series as a whole is unknowable, the procedure for computing every single term of it can be prescribed categorically in the form of a rule of general validity, stated once and for all, as in the case of integers.[5] Similar is the case with inductive proofs which are (infinitely) recursive by design, replacing endless case-by-case verification by one general logical argument. In contrast, the infinitesimalisation of a quantity involves division by an increasing sequence of numbers tending to infinity, i.e., arithmetic with unknowable – not just temporarily unknown (as in Bhaskara II's algebra, see below) – numbers. Jyeshthadeva's way out of this philosophical impasse is to fix a large number which he takes as *parārddham* $= 10^{17}$ – the largest *named* power of 10 in his

[5] The reader may find it useful at this point to go back to the relevant portions of Chapters 11 and 12 (and Chapter 10.4 for Nilakantha's explication of how the sum of an infinite series is to be understood as a limit). The sine series as treated in *Yuktibhāṣā* is less recursive in its logic since it relies on the exact finite formula for $\sin\theta$ in terms of $\sin\theta/n$ for finite n before taking the $n \to \infty$ limit. The treatment of Shankara in *Yuktidīpikā* is strongly recursive.

list of powers – and describe all steps in terms of this fixed named number, and then to treat it as a representative large number which is to be made as large as is necessary for the segments to be as small as we please (among the words used to characterise the smallness of the segments are *atyantam* (endlessly) and *śūnyaprāyam* (of the nature of zero)). It does not seem to bother him that the increasing sequence of numbers beyond whatever is the last named number cannot have names. (Or perhaps it did: the writing becomes uncharacteristically obscure in this part of the book). The strong endorsement of nominalism in the opening chapter of *Yuktibhāṣā* has been overtaken by the exigencies of the mathematics as it developed in entirely new directions; having paid his dues to the prevailing philosophical stance, always a step behind practice, he lets the mathematics take him where it did.

We may indeed argue that the invention of calculus ended the influence of nominalism as a guiding principle in mathematics.

Unlike zero, infinity does not exist as a number and did not have a name in India; the closest one got to it verbally was as an adjective, *ananta*, endless.[6] As we have just seen, the oral paradigm had an impact on both concepts, but in very different ways. On the zero the influence was, mathematically speaking, not deeply significant; it just postponed its arrival on the scene, whether as a name or, to a later epoch, as a symbol. The idea of infinity had been much more elusive to grasp as its long history from the *Taittirīyasaṃhitā* to *Yuktibhāṣā* demonstrates; we can in fact say that a precise sense of it began to be absorbed into the mathematical consciousness only after Madhava and his followers discovered that its mastery was a precondition to the making of a solid foundation for infinitesimal calculus. It is nevertheless a commonly held and stated view that the zero is among the greatest (if not *the* greatest) contribution of India to mathematics. That is clearly something of an exaggeration. In any case, the zero was used, and in a symbolic written form as a place holder, by the Babylonians from before the time of Pingala.

Did the weight of the oral tradition also suppress the development of an early form of symbolic mathematical writing and its use in formulae and equations? The fact that the only early written text we have, the Bakhshali manuscript (very likely pre-Aryabhatan, see Chapter 6.5), has a rudimentary symbolic and schematic notation which then fell into disuse, at least in 'published' texts, might seem to suggest that it did. But there are so many mysteries around Bakhshali – what caused written mathematics to disappear?; why is there no mention of Bakhshali or of its calculational schemes in any later source text?; does it represent another less widely practised peripheral tradition that did not survive the Aryabhatan renewal except, occasionally, in training manuals?; was it an importation of an alien method of recording mathematics?; and

[6]In the heyday of the making of number lists, *asaṃkhyeya* (uncountable) sometimes turned up as the name of various high powers of 10. The word *pūrṇa*, meaning full or complete, that occurs in the Upanishadic literature as a metaphysical concept is sometimes sought to be given the sense of the mathematical infinity, unjustified in my view; numbers cannot be both *pūrṇa* and *asaṃkhyeya*.

so on – that it will be rash to try and base any definite opinion on it alone. The computational schemata of Bakhshali are in any case generally accompanied by verbal narrations of the same material; we do not even know which came first: was the schematic version meant to hold the narrative version in a more abstract and economical form or was the latter meant to annotate the former?

The Bakhshali manuscript also brought in the innovation of denoting the unknown quantity in an algebraic problem by a symbol which has no semantic significance, a dot. The unknown already had a name from earlier times, *yāvat tāvat* (meaning, loosely, 'whatever it is'). The use of the names of colours to denote more than one unknown seems to have been first suggested by Brahmagupta. Later, Bhaskara II abbreviated the names to their first syllables and gave this notational device much theoretical prominence by placing it in the very first chapters of his *Bījagaṇita*. In the third verse of the book he says that fixed numbers and unknowns are to be distinguished one from the other by actually writing (*upalakṣaṇārthaṃ lekhyāni*) syllables for the latter and that negatives, unknowns included, are to be distinguished by marking a dot above. One would have thought that after this advocacy of symbolic notation (in unmistakable association with writing), Bhaskara would have illustrated its merits in his own monograph. Disappointingly, there is not one proposition or illustrative example (of which there are plenty in the book) that is stated or solved symbolically though there are allusions to equations and their manipulation; it is all words.[7] An explanation may be that *Bījagaṇita*, like many canonical texts, was intended to be the formal, public account of a body of knowledge most of which circulated among disciples and adepts informally. The reference to a written notation must mean that informal written notes were a normal part of teaching and learning. But that only deepens the mystery: why has no manuscript of an explanatory account of *Bījagaṇita* or *Līlāvatī* ever turned up (to my knowledge) that utilises mathematical notation?

A more fundamental explanation may lie in the place assigned to algebra, the true home of structural and abstract thinking, in the overall scheme of Indian mathematics. Bhaskara praises the virtues of the method of hidden numbers (*avyakta gaṇita* or *bīja gaṇita*) on various occasions but it is his pronouncement on what algebra is for, directly after the invocatory verse, that is revealing and relevant here: it is for the benefit of the slow-witted who may find the arithmetic of his earlier work (*Līlāvatī*) hard to follow. Stated otherwise, he saw it more or less as a computational short-cut and the unknown as a number like any other (which happens to be temporarily hidden); algebra was just a technique for making it *vyakta*, an adjunct to arithmetic. The idea of an object being defined by its structural properties but having no fixed value – which

[7] The problem immediately following the verse about the use of colour-syllables to denote unknowns is: "What is the result of adding one more than a positive unknown and eight less than twice the positive unknown?" The unknown is called descriptively *avyakta* (for *avyakta rāśi*), whose translation in the context is 'hidden number'; it is often used interchangeably with *yāvat tāvat*. Occasionally, the full names of colours stand for unknowns but not their recommended one-syllable abbreviations *kā(laka)* (for black), *nī(laka)* (for blue), etc.

is the source of the "generosity of algebra" (to quote D'Alembert) – is hard to find in Indian writing till we come to *Yuktibhāṣā*'s tutorial on polynomials (Chapter 13.2). There is in fact no standard word that is synonymous with 'variable', standing for any one of a whole set of elements sharing a common set of properties and/or subject to a common set of rules; that is the reason why I have mostly used the word 'unknown' rather than 'variable' in my references to pre-*Yuktibhāṣā* algebra, a distinction we do not always respect in current writing on the subject.

The contrast of course is with grammar. For Panini, language acquired structure and clarity by virtue of the self-consistent rules regulating its grammar, "rules without meaning" in Frits Staal's apt phrase (but not without purpose), rather than through the literal meanings of its words. His object of study, the grammar of a living language, was much more complex in its totality than the mathematics we have been looking at and the analytic techniques called for, consequently, intricate and demanding. It would not have been difficult for the best of the mathematicians – who almost surely got a grounding in the grammar of Sanskrit during their formative years – to appropriately streamline and adapt the Paninian paradigm with its rigorous framework of principles and precise rules of implementation to the simpler needs of mathematics. That happened only to a very limited extent. The influence of grammar was largely confined to routine matters like the correct construction of a technical vocabulary; in the mathematical universe, the Word dominated over the Rule, Brhaspati-Indra enumeration (with help from Bhartrhari?) over Panini-Patanjali abstraction.[8] It is strange and, to a present day sensibility, astonishing that it should be so. Almost all of Panini's formal machinery could have been adopted by the mathematicians – as indeed it has been in modern times – with profit: from metalinguistic principles of organisation (formal set theory without the axioms) to the metatextual style of presentation, intelligible only to the initiated, but unerringly precise.

It is futile but fascinating nevertheless to ponder how a whole-hearted adoption of Paninian structural methods might have transformed India's mathematical landscape.

14.3 Geometry with Indian Characteristics

Knowledge of early Indian geometry reached Europe (and a few Euoropean Indologists in colonial India) after Thibaut's translation of Baudhayana's *Śulbasūtra* was published in the second half of the 19th century. It was a pivotal period for European mathematics; Europe had had three centuries to assimilate

[8]Pingala was an exception but he was ignored by mathematicians until very late, thought of probably as a prosodist. Aryabhata's short-lived *makhi, bhakhi, . . .* code for names for numbers was just a clever isolated trick which did not respect the decimal place value principle, devised to enable him to fit the large numbers pertaining to planetary parameters into a few verses.

Euclid's geometry, not only the beautiful theorems in the *Elements*, but, in a more profound sense, also the method of deductive reasoning that produced these theorems from a handful of postulates or axioms (and rules of reasoning), supposedly self-evident, once the objects to which they were applied were properly defined. Already in the 17th century Euclid had been installed as the aspirational model for all scientific (e.g., Newton) and even non-scientific (e.g., Leibniz) enquiry, though not, even in mathematics, with a great deal of success. Indeed, what followed was the freewheeling and spectacularly productive mathematics of the 18th century and the first half of the 19th, a golden age that was too busy being successful to pay much attention to the Euclidean ideal and its axiomatic rigour. That did not last; by mid-19th century, these very same advances had made it impossible to ignore the poorly understood foundations on which many of them were based. The objects that mathematics dealt with and the rules that governed their manipulation came under intense scrutiny including, among the subjects we have touched upon in this book, the recursive structure of (positive) integers, the approach to the infinitesimal through limits of infinite sequences, a proper characterisation of irrationals (also as limits of sequences) etc. Formal logic as the touchstone of correct reasoning was elevated to the status of a (meta) mathematical discipline in its own right, going far beyond Aristotelian syllogisms.

It was in the midst of this ferment that European (and, only a little later, American) Indologists as well as some fine mathematician-historians first came across the collection of geometrical results of the *Śulbasūtra*, many of them familiar from Greek geometry but none with an accompanying proof. Coming from a part of the world whose mathematical culture was largely a closed book,[9] they struck a responsive chord among those who had already embraced enthusiastically William Jones' hypothetical Indo-European as the common ancestor of European and Indian languages and quickly found a place in one or two accounts of mathematical history, notably in Moritz Cantor's encyclopaedic work. Attempts were made to reconstruct plausible demonstrations and, as far as purely geometrical results such as the Pythagorean theorem were concerned, they converged to the view that the Vedic seers indeed had proofs, generally in a cut-and-fit style described in later Indian writings (see Chapter 2.2). Some European scholars also dated the earliest *Śulbasūtra* (Baudhayana and Apastamba), to a time prior to Pythagoras, mainly from philological considerations. That conclusion has stood the test of time. (It helped that Pythagoras and Panini were rough contemporaries).

Thibaut did his work in India and he had the acknowledged help of traditional Indian experts in the *śāstra*. The reaction of the professional British

[9]The much earlier publication of Colebrooke's translation of Bhaskara II and Brahmagupta and the work that preceded it seem to have influenced only Playfair from among those who had the necessary background. No one else wrote about Aryabhata's trigonometry (the Kern edition of the *Āryabhaṭīya* was published only in the mid-1870s, just a year or two before Thibaut's translation of Baudhayana) and none at all seems to have followed up on Whish's London lecture.

indologists living and working in India (one of them being Kaye whom we have met in connection with the Bakhshali manuscript) to the same *Śulbasūtra* geometry was markedly different from his own and from that of continental Europe. Their real concern was to ensure the primacy of Euclid in terms of content and timing. It was a given for them that Indian mathematics was derivative; geometry in particular was Euclid reformulated badly or, even if the *Śulbasūtra* had existed in some primitive pre-Euclid form, the interesting bits in them were all later insertions. The late 19th and early 20th centuries were not a time of glory for British scholars of India's scientific past even as pioneering work was being done by text translators and archaeologists. Comparisons with Euclid were in any case natural and inevitable; the more elementary portions of his geometry had by then become part of the Classical inheritance of every educated European. With the benefit of hindsight, we can see now that no one with even a vague feeling for the two respective geometric cultures could have taken such a position and that these aberrant views were bound to fade away, as they did. Questions of priority were to surface again later around other mathematical topics (with less excuse it may be added), and not only in British India (see Chapters 7.4 and 12.4 for instances; also see below).

So, after this recapitulation of generally accepted historiography, what are the hallmarks of geometry with Indian characteristics? To begin with its wellsprings, the first and lasting impression is that the *Śulbasūtra* are collections of assorted results of varying degrees of generality, exactness and purity (in the sense in which 'pure' is used in modern mathematics), organised somewhat haphazardly. The later ones of Manava and Katyayana take little care about the logical interconnectedness or sequencing of the material they present and, with one or two interesting exceptions, there is little in them that is not already found in Baudhayana and Apastamba. So we may assume that the *Śulbasūtra* taken together is a fair portrait of the state of geometric knowledge in the first millennium BCE (and that it did not change much during that time), with a few elements harking back all the way to the architecture and decorative arts of the urban Indus Valley people. The apparent disorganisation probably reflects an organic development in response to architectural demands. Or, since no proofs were intended to be given in the *Śulbasūtra* – and, possibly, none known to the later architects of *vedi*s, rather as most practising astrologers in modern Kerala have forgotten the astronomical basis of their astrology – the need to explain the causal and rational connections among the propositions was probably not felt. Altogether, it is a far cry from the completeness and organisational perfection of Euclid for whom each proposition followed relentlessly from those already established. And there was no authoritative philosophy of knowledge in India before, say, the 6th century BCE to guide the architect-geometer securely from cause to consequence – that came later, in several variants – a handicap Euclid certainly did not suffer from.

With the rebirth of geometry in the Aryabhatan era, mathematical texts as a whole and geometric ones in particular became substantially better organised, as we can see in *Āryabhaṭīya* and *Brāhmasphuṭasiddhānta*. What did not

change much was a certain quality of visual communication that Indian geometry always possessed. Diagrams as reconstructed from texts – they must have been used – often convey directly the essential geometric truths behind them. They relate to real space in which real objects can be constructed, they are not just props supporting a chain of disembodied logical argumentation – compare the cut-and-fit proofs of the diagonal theorem (Chapter 2.2) with the artificial constructions involved in Euclid's proof. The longevity of such 'organic' methods – another example is Nilakantha's 2- and 3-dimensional demonstrations of Aryabhata's series identities (Chapter 7.1) – is the best testament to the continuing popularity of this visual geometry. There are also other, less immediately apparent, instances of its efficacy and appeal: the diagram of Aryabhata's half-chord as a side of the right triangle associated naturally to an arc (Figure 7.3), the progression from the pre-Brahmagupta to the Brahmagupta quadrilateral (Figures 8.1 and 8.2) and Narayana's third diagonal (Figure 8.4) are visual paraphrases of nontrivial geometrical truths. It is pertinent to note that intelligent visual inspection and cut-and-paste procedures never led to conclusions which were wrong and that all the propositions purportedly arrived at by such means can also be established by strictly Euclidean methods once the relevant objects – circles, lines, squares, rectangles, etc. together with the notions of length and area – are given explicit definitions; and that is an instructive thing to do.

The quintessential geometric object in all these is the circle (which occurs relatively late, for the first time in Book III, in Euclid) and the quintessential geometric property is the orthogonality of a chord and the diameter bisecting it (Indian orthogonality as I have called it earlier), together with the theorem of the diagonal. We have already noted that, as a consequence, the most fundamental relationship two lines can have in India is that of orthogonality, not parallelism, so much so that the term *tribhuja* (trilateral) without qualification came to denote almost exclusively a right triangle. The circle and the right triangle consequently became the main ingredients in the continuity of the mathematical tradition as it applies to geometry, as we have seen repeatedly in the course of this book. Nothing exemplifies their ubiquitous and permanent presence better than the more or less contemporaneous development of trigonometry and the theory of cyclic quadrilaterals both of which, also as we have seen, can be traced straight back to a common source 1500 years in the past, the *Śulbasūtra*. Indeed, all of Brahmagupta's work on real and integral cyclic quadrilaterals is a direct outgrowth of the diagonal theorem and the existence of rational diagonal triples and their scaling property (Chapter 8.4). Mathematical themes never became obsolete in India and one may well think of the evolution of cyclic geometry as the perfect illustration of a certain innate conservative streak in Indian mathematics, its sedate progress (or, often, no progress at all over long periods) punctuated by episodes of dazzling creativity. Narayana's introduction of the third diagonal, 700 years after Brahmagupta, brought this particular theme to a close. That these episodes came at such widely separated times is also very Indian and only strengthens the impression

of conservatism: compare the chronology of the golden age of Greek geometry, say from the Pythagoreans through Euclid to Archimedes and Apollonius, which took a mere 300 years or so.

To an extent, it is unavoidable that any attempt to identify what is Indian about Indian geometry should begin with a comparison with the geometry of Euclid; so overwhelming has been its influence that, subconsciously at least, geometry is Euclidean geometry for most of us. There is another reason why it deserves our attention: the Indian geometric tradition has the backing of authentic and fairly complete textual resources even if nowhere near as comprehensive or easy to follow as Euclid,[10] making a comparative textual analysis possible. In this spirit, I bring up two other points of comparison, both of which have been discussed by historians but inadequately in my view: the question of proofs and the extent to which numbers and arithmetic were (or were not) involved in the geometry. It is more practical to get the second out of the way – there is plenty of explicit textual material to draw upon, much of which we have already seen in the first three parts of this book – before turning to the all important matter of what the idea of a proof meant to the Indian mathematician. There is enough to be said on this subject and on proofs in general, not just in geometry, to justify devoting the whole of the next chapter to it.

It is an old and widely accepted observation that Indian geometry is less pure than that of Euclid (and other Greeks who lived before astronomy came into the picture in a big way) if by 'pure' is meant 'not tainted by numbers'. The great majority of Euclid's (geometric) propositions have no reference to specific values of lengths etc., limiting themselves at best to statements of comparison: equal, bigger or smaller. Nothing much in them will have to be rewritten if an assumed numerical value of say, π, which was not very precisely known in Euclid's time, were to change, as it of course did. The theorem of Pythagoras is a proposition like any other, with not too many applications made of it. In India on the other hand, it was the key to the kingdom of geometry: the metric was always part of geo-metry, which was as much about measurable magnitudes as about shapes: the abstract diagonal theorem in the *Śulbasūtra* is closely accompanied by lists of integral triples, the numerical values of π and $\sqrt{2}$ are an inseparable part of the geometry of the circle and the square, and so on. One explanation for this easy cohabitation is the essential role of scaling in the architecture of the *vedi*s since their absolute dimensions had to be in proportion to the 'size' of the patron of the sacrifice (the *yajamāna*). A deeper reason – and one which made scaling itself possible – was that the growth of geometric knowledge was accompanied or preceded by the mastery of the decimal number system. There was no fear of numbers and, as a corollary, approximate geometry of the kind represented by the area-equivalent construction of squares and circles was not cast out of the fold of geometry. An unqualified acceptance

[10] A detailed decoding of Babylonian geometric tablets remains largely a task for the future. Some of them may have interesting things to tell us about the Babylonian rationale for the theorem of the diagonal.

of numerically approximate geometry was epistemically crucial in many later developments such as trigonometry.

But why is an explanation felt to be needed at all for such a natural union, the weaving together, so to say, of the warp and weft of the very fabric of mathematics? On the contrary, one may well argue, and I do, that what needs to be understood is the Euclidean separation between numbers and space. Greek mathematics at the time of Euclid (and later) had no systematic rule-based method of handling numbers in their totality. Lacking the idea of a base, Euclid could not have treated arithmetic with anything like the admirable rigour he brought to qualitative geometry[11] and perhaps he did well not to try to mix them. It is a tantalising thought that the great emphasis put on ratios and proportions – in other words measures of deviations from equality – and, especially, the many inequality propositions in his geometry, are themselves a reflection or a residue of the comparison-counting or matching of the kind that, in India, came before based decimal numbers (Chapter 5.1). The term 'magnitude' is everywhere (though itself undefined geometrically) in Book V (on Proportion) and its long list of definitions, but the magnitude is always illustrated by the (unspecified) length of a line; there is no concrete number, it is all comparison. (The only actual number I have spotted in the first five Books is 2 in, for example, the proposition which says that the area of a parallelogram is twice that of a triangle with the same base and altitude). Much has been written in modern times about how these cogitations on the comparison between (geometric) magnitudes, generally taken to be inspired by Eudoxus, laid the foundations for our understanding of the notion of a limit, which is true though it took 2000 years starting with Archimedes (see Chapter 12.4) to work it out. But at least the writing of Book V could have been made easier if the Greeks had absorbed the conceptual significance of a base from the Babylonians. Even its practical utility was poorly exploited; well after the sexagesimal system became known to them, astronomers (who could not possibly escape from quantitative geometry) in the Hellenic world still had difficulties in integrating numbers into their geometry, as we saw in connection with Ptolemy's chord table (Chapter 7.3).

To conclude this brief foray into the early history of the interface between 'pure' geometry and numbers, let us recall that the concept itself of a distance (in spaces of dimension greater than 1) depends on the diagonal theorem, directly in 'flat' spaces, and through its infinitesimal form when the space has curvature. The latter recognition also we owe to the Nila mathematicians, specifically to *Yuktibhāṣā* (the effect of curvature on calculations based on the rule of three (linearity), see in particular the quotations in Chapter 7.5 and Chapter 10.3). Geometry has become infinitely richer in every sense since then, but the indispensability of the diagonal theorem in measuring space remains: Cartesian ('coordinate') geometry, which finally brought numbers and space together in Europe, would not have been possible without it.

---

[11]Even if he had an axiom system for numbers, it would not have helped. No one does practical calculations (*vyavahāram* in *Yuktibhāṣā*) using Peano's axioms; one needs a convenient base.

15

What is Indian . . .? The Question of Proofs

15.1 What is a Proof?

Just as the search for what is Indian in Indian geometry turned quickly into a comparison with Euclid, so is an inquiry into what guarantees the correctness of general mathematical statements – how did Indian mathematicians *know* that their propositions were 'true'? – best initiated through geometry, and for the same reason: we have in Euclid's deductive system an unambiguous standard, within a certain epistemic framework, with which comparisons can be made. But, before that, it is helpful to note that the question has begun to attract attention only relatively recently, initially among Western scholars, and mostly – so far – through casual remarks not backed by careful analysis. Indian scholars of early modern times did not think there was an issue, perhaps naturally; nor did the Thibauts, the Rodets and the Cantors and several others of the late 19th century. The reasons for the European (and early 20th century American) indifference to the question are easy to imagine: in the first flush of the discovery of a whole body of exotic geometric knowledge – Pythagoras' theorem, stated precisely in words with no decoding or guesswork needed, three centuries before Pythagoras! – the problem of how the cut-and-move proofs they came up with could be made to follow from Euclid's postulates and common notions was probably the last thing on their minds. And, after two exhilarating centuries of logically unfettered mathematics, axiomatism itself was not as strong a presence as it soon became: think of the rich and beautiful number theory as it existed before Peano got around to telling us what numbers are.

Whatever the reason, the realisation soon came that the epistemic foundation of Indian mathematics, geometry to begin with, was very different from

P. P. Divakaran, *The Mathematics of India*, Sources and Studies in the History of Mathematics and Physical Sciences, https://doi.org/10.1007/978-981-13-1774-3_15

that of Euclid.[1] At a superficial textual level, these differences are easy to spot. No Indian work has a list of postulates, in fact there are no postulates; the category of metamathematical principles which served the same purpose is unspecified but, we must suppose, tacitly agreed to by all, forming a common pool of accepted facts. One such fact, frequently used but never stated, could be: two lines in the plane, both perpendicular to a given line, will not intersect, no matter how far extended. (If Euclid were to have given priority to orthogonality, he might have stated the parallel postulate thus). Definitions are not spelled out. No one, not even the meticulously logical Jyeshthadeva, bothers to define the circle, drawn with a pair of cord-compasses, as the locus of points equidistant from a given point, nor are we told what a point or a line is. For the Indian mathematician, it would have been as pointless to pin down the notion of a point (in terms of what?) as it would have been for the grammarian to define a syllable or for the logician to define Devadatta as the generic subject in syllogistic statements; everyone knew what they were. When it came to less 'atomic' concepts, the technical vocabulary was supposed to take care of the need to be precise. Generally speaking, it did, but we have also seen instances where it is incomplete or not sufficiently specific: *jyā* (= chord) for *ardhajyā* (= sine = half the chord of twice the arc) and even for *ardhajyāntara* (= sine difference); *viṣama* (= unequal, the opposite of *sama*) for a general cyclic quadrilateral, etc., etc.

These deviations from what we consider good practice today are to be found in texts ranging over all periods and covering all branches of mathematics and so must be considered generic.[2] So must some other equally obvious deficiencies: objects to which a statement applies are often not specified, conditions under which a proposition holds are not stated; special cases and approximations are not always identified as such, and so on. Sloppiness of this sort did not by itself cause any serious misunderstandings, two notorious instances excepted, both concerned with geometry, both from the same verse of *Brāhmasphuṭasiddhānta* on the area of a cyclic quadrilateral (and both discussed in Chapter 8.4): *tricaturbhuja* misread as triangle *and* quadrilateral by almost all commentators ancient or modern (including Sarasvati Amma in her book [SA]) and *viṣama* whose correct reading as cyclic quadrilateral was missed by no less a master than Bhaskara II. That the mathematical substance

---

[1]As I said above, serious analytic studies are still in their infancy; readers may find it useful to consult the writings of Roddam Narasimha of which the most easily accessible is "Epistemology and Language in Indian Astronomy and Mathematics", J. Indian Philosophy, vol. **35** (2007) p. 521 (for the philosophy) and M. D. Srinivas' Epilogue "Proofs in Indian Mathematics" to K. V. Sarma's English edition of *Yuktibhāṣā* (Part I) [YB-S] (for a few examples).

[2]There are two partial exceptions to this observation, *Yuktibhāṣā* and Nilakantha's *Āryabhaṭīyabhāṣya*, both of which are very late texts in prose. We shall see presently that they also mark a turn towards more formally structured proofs. *Yuktibhāṣā* presents an interesting compromise; while it is careful in describing the new notions arising in infinitesimal mathematics and in constructing a satisfactory mathematical vocabulary to serve its ends, it sometimes lapses into conventional and literally inaccurate terminology when invoking the mathematics of the past. Force of habit perhaps.

survived this lack of concern for clarity of writing and terminology was surely a tribute to the robustness of the oral tradition; the initiated knew what every term stood for and how a sentence was to be parsed, a knowledge that outsiders like us, the historian and the reader, have to earn.

For outsiders, to go from the visible evidence of clumsy terminology to the conclusion that it reflects a more fundamental defect, that of a lack of precision in the underlying concepts, is not unnatural. Equally visibly, original texts, from the *Śulbasūtra* to Narayana's *Gaṇitakaumudī*, do not as a rule provide any sort of justification for their assertions. Commentators (among whom we should include Bhaskara II) are a little more explicit with their explanations and interpretations of difficult points, but will disappoint those looking for complete and well laid out proofs in the modern manner. From this combination of lapses from Euclid's (and today's post-Euclidean) standards of disciplined discourse was born the widely shared belief that Indian mathematicians had no proofs for their discoveries or, worse, did not have logical criteria for what a proof should be. It is now clear – though apparently still not to everyone – that this is a myth whose parentage goes back to the early 20th century British orientalists. First of all, the texts of the great masters, especially those in the *sūtra* format, were meant to be a permanent record of knowledge newly gained, not to justify it; that was done in the parallel pedagogic activity of face-to-face instruction and debate (see below for Nilakantha's views on the validation of new knowledge). Then one must ask how, to quote Aryabhata, the best of gems that is true knowledge are to be brought up from the ocean of true and false knowledge if not by the boat of one's intelligence; surely not by divine inspiration or divination. And not by guesswork alone either: it is a telling fact that, in the vast treasure house of India's mathematical "gems", there is very very little "false knowledge" (ironical exceptions: two volume formulae of Aryabhata's own) and none at all that relates to anything deep and original. That is an impossibly high rate of success for mere conjectures, even by the standards of modern mathematics. It would be naive indeed to think that it all came out of anything other than the exercise of one's own intelligence, *svamati*, within a consciously erected theoretical frame; they must have been doing something right.

From today's vantage point, the question of what their conception of a proof was is best answered by going through the actual proofs in *Yuktibhāṣā* which, as I have already said several times, supplies a *yukti* to virtually every one of its many propositions. (I repeat that the accounts of the more important of these proofs in Part III of this book are as faithful to the original as is consistent with modern notation and with my aim of making comparisons with later European work easy). *Yuktibhāṣā*, being in Malayalam, was inaccessible to most scholars until the English translation was published recently and it is understandable that it had no influence on the early discussions of the status of proofs. That cannot be said of Nilakantha's *Āryabhaṭīyabhāṣya* which was published in the mid 1930s, is in Sanskrit and has crystal clear derivations of all the results of the *Gaṇita* chapter. And it is not as though the earlier

literature was devoid of any mention of *upapatti*, a term, used from the 7th century onwards at least, having a whole spectrum of meanings: evidence, cause, reason, rationale (especially in the context of mathematics), etc. The imperative need to have an *upapatti* before a mathematical statement can be taken to be true is emphasised more than once by Bhaskara II (and others) and proofs of varying degrees of completeness and structural clarity do occur in many pre-Nilakantha commentaries. An instructive example is the formulae of Bhaskara II for the surface area and volume of a sphere: we saw in Chapter 7.5 that he starts promisingly (finite integration) and gets the right results but is unable fully to justify them, understandably because that would have required taking the infinitesimal limit. One cannot doubt that the essential idea of a proof is there, though it may not be an easy job to isolate its structure from the general narrative and complete it – in this particular case, by Jyeshthadeva. Similarly incomplete is Bhaskara's description of the method of *cakravāla* which got a clear exposition only in the work of modern commentators, as far as I can tell. (Other examples can be found in Srinivas, cited above). In spite of such minor reservations, and even setting aside the writings of the post-Nilakantha Nila phase, the notion of Indian mathematicians not really knowing what they were doing, and why, can have only one resting place: the place Aryabhata calls the ocean of false knowledge.

As a general rule practising mathematicians in India, like those elsewhere, did not mix their science with its metaphysical and logical foundations, with one fortunate exception. When someone as articulate and as deeply reflective about his vocation as Nilakantha speaks about how we acquire new knowledge and how we are to know it to be true (the three wise books, Chapter 9.2), we have to listen. It is more than likely that his thoughts are not just personal but mirror the epistemic ethos of his time and place and perhaps that of Indian mathematics as a whole from the time of Aryabhata; after a thousand years, the insistence on the supremacy of the intellect (*mati*) continues to echo in these pronouncements. In slightly oversimplified terms, here is the model he proposes (a more analytic account in the canonical philosophical language will be found in Narasimha, cited above, and references therein).

The primary instrumentality of our apprehension of the world is our senses, not only of the astronomical world but even of the mathematical (visual geometry such as in cut-and-shift proofs, 3-dimensional geometric versions of algebraic identities, etc.). The sense data are to be subjected to analysis and tentative inferences drawn by means of our mental faculties, our ability to compute included. These are then exposed to scrutiny by the knowledgeable, debated and revisions made if necessary, and shared with pupils, thus sustaining a living chain of continuity. Nilakantha would not go out of his way to run down revealed wisdom, *śruti*, (his remote predecessors from 9th century Mahodayapuram in Kerala had already done that with their derisive remarks about *paurāṇika śruti*, the supposed revelations of the Puranas) and he has reverence for the words of the great teachers from the past. But he is uncompromising

about the need to subject prior knowledge, whether revealed or merely uttered by mortals and lodged in an abstract communal memory (*smṛti*), to the tests of observation and logical inference and rejected if found wanting. In an uncanny paraphrase of Aryabhata's "*svamatināvā*", but more modestly, he says: "Everything here (the proof of the theorem of the diagonal) is rooted in *yukti* (about which, see below) alone, not in the beliefs and practices of yore (*āgama*)". And, as for the primacy of observation, he practised what he preached, taking upon himself the mission of campaigning for his teacher Parameshvara's *Dṛggaṇita* revision, forced by fresh data, of the planetary model of Aryabhata, no less.

The principles on which the practice of mathematics in India was grounded were not, thus, very rigid or 'theoretical' (Nilakantha says that theories are unending and inconclusive) and they could not have been more different from the inflexible frame in which Euclid's austere axiomatic-deductive system was confined. There were no unquestionable first principles to help choose, once and for all, a set of postulates and rules of logic, or to decide what objects needed to be defined. Instead we have a more dynamic and fluid foundation that was built from intelligent, rigorously exercised common sense so to say, and responsive to accumulated experience. The idea of an infallible set of axioms leading, by equally infallible logic, to mathematical truths of unquestionable certitude would not have been given a hearing: if all knowledge is contingent, how can it be otherwise for metaknowledge, the knowledge that some part of that knowledge is (or is not) true?

How then did Indian mathematicians know or decide that their mathematical insights were indeed true? The answer, the only satisfactory one given their philosophical stance, is in the last two of Nilakantha's tests of validation: there is no absolute criterion but only a convergence towards a consensus; the search for the infallible proof is ultimately a futile search, "unending and inconclusive" like all theorising. In practice most mathematicians seem to have been happy to leave the judgement to their peers, as Bhaskara II suggests in his comparison of a putative proposition whose *upapatti* does not get the approval of the assembly of the learned to rice without butter, unpalatable. The miracle is that the mathematics that this philosophical openness produced is true (and interesting) mathematics by contemporary axiomatic-deductive standards. Or, perhaps, there is no miracle; perhaps the universal and immutable truths of mathematics are open to all gifted and prepared minds – even of those who never heard of Plato and Aristotle.

Nilakantha's musings on the nature of mathematical truth and how to pin it down go hand in hand with the way proofs are laid out in his own commentary on the *Āryabhaṭīya*, its *Gaṇita* chapter in particular. There is much emphasis on the organisation of the reasoning employed, approximations are avoided when there is no need for them – as in his proof of the sine difference formula and Aryabhata's rule (see Chapter 7.4) – the sequencing of steps is carefully causal and conclusions are drawn almost painlessly. We can see the same insistence on structural clarity in proofs, but on a broader canvas, in *Yuktibhāṣā* which, as I have argued earlier, is of the same date. (To appreciate what a big change this

represents, it is enough to compare with Bhaskara II's writing on the topics they both treat, cyclic quadrilaterals or the sphere formulae for example). The importance of laying out the logical structure of a complex line of reasoning, as part of the obligation of a mathematician-preceptor – we no longer have to second-guess or 'interpret' the text – was thus finally given its due.

As if to advertise these newly recognised virtues, there is a subtle shift in the terminology as well, the old *upapatti* making way gradually for the term *yukti*. Nilakantha uses it freely and Jyeshthadeva likewise, in addition of course to giving his book a title that highlights its special place in the new discourse. The word *yukti*, cognate with *yoga*, literally means 'joining together'. The sense which the term is given by Nilakantha and Jyeshthadeva is consistent with this meaning (it has a whole collection of related auxiliary meanings): a chain of logically sequential steps, well laid out, generally not short, that ends in a significant proposition that was far from obvious to begin with. (The word *upapatti* does occur here and there but for simpe inferences). For our purpose, that of a quick look at the logical basis of Indian proofs, it seems perfectly acceptable to translate *yukti* simply and functionally as 'proof', as long as it is kept in mind that the logic was very different from what is axiomatic for modern mainstream mathematics; we shall look at one especially dramatic manifestation of the difference in the next section.

15.2 No Reductio ad Absurdum

Nilakantha's scepticism on the admissibilty of an absolutist position on meta-mathematical questions has tangled roots in the history of Indian philosophy, going back to the late Books of the *Ṛgveda* (See Chapter 6.4 on Aryabhata's invocation of his only god (or non-god) Ka). The agnosticism continued into the Upanishadic times, in the face of the strong currents of ritualism that ran through middle Vedic Hinduism (a convenient term even though it can be given a historically meaningful sense only in the light of the new religions which were, at that time, still in the future). Buddhism and Jainism turned their backs on visionary speculations and mechanical rituals alike, and ushered in the idea of rationality as the unifying principle of all discourse: the world in all its complexity, from individual life to the fullness of the cosmos is to be understood through the agency of one's intellect. The willingness to acknowledge that everything cannot be known, however, remained. The Buddhist canon has passages in which the Buddha admits the limitations of rational enquiry regarding matters not only of faith but also the material universe: for example, does it have a beginning and an end in time and space?

Fierce disputes ensued and along with them came an interest in the mechanics of disputations: how can one judge the veracity of a declaration and what are the legitimate means of buttressing or refuting it? If a date is to be assigned to the genesis of logic as a discipline on its own, this period, loosely 6th-5th century BCE, would be a fair choice for India. The Buddha himself

believed that logical argumentation (sometimes tediously long) can be applied usefully to all questions of interest to man, mainly through an analysis of causation. At the same time, he rejected what later became one of the pillars of Aristotelian logic (and Western logic, more generally, ever since), the notion that logic is bivalent, that, in the context of epistemology, every sensible proposition is either true or not true[3] (or, synonymously, false; it is not a bad thing to be pedantic as the subject is not free from semantic ambiguity).[4] The evolution of Buddhism into its Mahayana form in India did nothing to dilute its suspicion of bivalence. It in fact crystallised the distrust in a formal structure, Nāgārjuna's so-called tetralemma: the four classes {true, false, true-and-false, neither-true-nor-false} exhaust all logical possibilities regarding any sensible (non-nonsensical) statement (semantics again, and this is not the first time nor the last that language will impinge on logic).[5]

Outside Buddhism also, the study of models of logic flourished, influencing heavily at least one, *Nyāya*, of the conventional six schools of philosophy. The heyday of *Nyāya* overlapped with the general descent into Puranic irrationality in the centuries following Aryabhata and may well have been a reaction to the myths and magic of the times.[6] Building on the work of their Buddhist predecessors and contemporaries, some of the most prominent teachers of the school paid special attention to the problem of logical negation and its semantic expression, in particular the validity or otherwise of the proposition according to which double negation is equivalent to affirmation (which, elevated or relegated to the status of an axiom, is at the core of bivalence).[7] Partly because of the

[3]There are several places in the canonical literature in which the Buddha rejects the dual complementarity between 'is' and 'is not'; the 'middle way' (*mādhyamika*) connotes, in one sense, equidistance from the two extremes. My guide to the subtleties of the Buddha's dialectical stance is T. R. V. Murti's classic *The Central Philosophy of Buddhism* (2016, reprint), Motilal Banrsidass, Delhi, especially the first four chapters.

[4]Modern logicians have several technical terms for several variants of approximately the same idea and having approximately the same uses. They are hardly relevant to the understanding of the Indian views on the subject of mathematical truth. In what follows, 'bivalence' can be taken to be the same as the 'rule of the excluded middle' as far as mathematical statements are concerned, see below; similarly, I use the term 'proof by contradiction' interchangeably with 'proof by *reductio ad absurdum*'.

[5]The first intimations of logical/linguistic tetravalence are already present in the *Ṛgveda*; Book 10 speaks of the "Uncreate", the undefinable pre-creation state, as "neither non-being nor being" and, simultaneously, "being and non-being".

[6]I note in passing that *Nyāya* too is one of those words with a wide and rich range of meanings, one of which is 'a universally valid rule', appropriate both in its association with logic and the sense in which it is used in mathematics: *bhujā-koṭi-karṇa-nyāya* (the theorem of the diagonal), *jīveparaspara-nyāya* (the addition theorem of sines), etc.

[7]Those who would like to delve more deeply into the subject will find invaluable the writings of B. K. Matilal, for instance the books *Logic, Language and Reality: Indian Philosophy and Contemporary Issues* (1985), Motilal Banarsidass, Delhi, and *Perception: an Essay on Classical Indian Theories of Knowledge* (1986), Oxford University Press, New Delhi. A reader-friendly general survey is Sundar Sarukkai's *Indian Philosophy and Philosophy of Science* (2005), Centre for Studies in Civilizations, New Delhi. These books will at least serve to remove the misconception some historians of mathematics had and still have that Indian thinkers were not, to put it baldly, concerned about epistemic issues.

absence of a notation that could bypass common linguistic usages and even more because the universe of the discourse was not delimited sufficiently sharply (the two are not unrelated, see below), no consensus emerged that might have been of interest in mathematics. Argumentation by contradiction even had a name, *tarka*. But at the end of all the fine-spun philosophy, there seems to have been a tacit but unanimous collective judgement in mathematical circles against proofs by contradiction: there are no such proofs in Indian mathematics.

Leaving metaphysical speculation aside, formulating a logical criterion for what constitutes a proof of a mathematical proposition would seem to be difficult if not impossible without the anchorage provided by a metamathematical axiom system. (Whether that is enough is a different matter as is the continuing debate among axiomatists about the merits of this or that set of axioms). At the very least, it can help to minimise ambiguity in the unavoidable verbal communication that is essential to the creation of the body of shared knowledge that we call mathematics. Indian logicians struggled with the innate imprecision of language all through (as did the Greeks: Aristotle had to hedge his statement of, say, the law of non-contradiction with phrases like "in the same sense"; what does it mean, objectively?). As far as logical bivalence is concerned, the prime source of ambiguity is the word 'not': what colour, to an Indian logician (the example occurs), is a flower that is not not-white? To paraphrase Patanjali, it seems that some general rules of logic must be set down before the particularities of how to sharpen the linguistic 'not' to the logical **not** are tackled. That task, in turn, seems best approached in the language of sets, something the *Nyāya* philosophers came close to doing without actually getting there.[8]

The indispensable general rule is the specification of a universal set U consisting of all the objects about which statements are intended to be made. That means axioms or rules imposed on operations on U, for instance U could be the set of non-negative integers as defined by the operation of succession along with the (Peano) axioms it is subject to. The odd positive integers then form a subset S of U defined by enumeration skipping one in the manner of the *Taittirīyasaṃhitā*: $S = \{1, 3, . . ., 99, . . .\}$ and the set of non-negative integers which are **not** odd is well-defined as the complement of S in U: $S' = \{0, 2, . . ., 100, . . .\}$ with the property of bivalence, $(S')' = S$, satisfied. Without the general rule specifying U, the particular rule implementing **not**, which is the formation of the complement S' given S, is ill-defined and therefore meaningless. The point is not so much that the specification of U is enough to define **not** as the enforcer of bivalence – we know from the recent history of formal logic that that may not always be the case – but that, without it, we may not even make a start; the hierarchical order of rules must be respected.

Staying within the (unIndian) axiomatic framework, the prototypical *reductio* proofs are proofs of irrationality and of them the simplest is for $\sqrt{2}$. I recall here its most elementary and widely familiar version because it is simple

[8]Given a property p, the subset of objects having it (the locus P of the property p) was considered in discussions of negation; see Matilal, *Logic, Language, ...* (cited above), pages 114 - 119; the set of which P is a subset is not, however, explicitly defined.

enough to illustrate, without other distractions, the structure of such proofs and so can help us get a feeling for the Indian dilemma regarding them as evidenced in Nilakantha's assertion of the irrationality of square roots and π without offering a shred of proof (Chapters 9.2 and 13.4). Choosing U to be the set of real numbers and S to be its subset of rational numbers, a number that is **not** rational gets defined as an element of the complement S' of S in U. By bivalence, $\sqrt{2}$ is either in S' or S. To prove that it is in S', one begins by supposing that it is in S, i.e., that it can be expressed as a fraction, $\sqrt{2} = a/b$, a and b (positive) integers. Suppose further that a/b is an irreducible fraction (a and b relatively prime) as, by the rules of arithmetic, dividing a and b by their highest common factor does not change the value of a/b. It follows that a^2 ($= 2b^2$) is even (**not** odd); hence a is even; hence a^2 is a multiple of 4; hence b^2 is even; hence b is even; i.e., a and b have 2 as a common factor, which contradicts the initial prescription on them. Therefore, by bivalence but applied this time to positive rationals as the universal set and irreducibility as the property on which **not** is defined, we arrive at the contradiction that a rational $\sqrt{2}$ is **not** irreducible if it is irreducible. Hence $\sqrt{2}$ is **not** rational.

From the viewpoint of its logical credentials, the first thing to be noted about this standard proof is that the logical **not** is invoked a number of times (some of which are implicit and may be thought to be trivial, e.g., in concluding that if a^2 is even, then a is even) and the contradiction established directly is not of the primary bivalence defined by rationality itself, but an auxiliary one defined on the set of positive rationals by ireducibility. In other words, the proof depends on standard arithmetic to the extent of the validity of the fundamental (unique factorisation) theorem of arithmetic without which irreducibility of fractions is a meaningless concept. The point, once again, is not whether unique factorisation of integers is questionable, but that the proof requires a sequence of steps – it will be historically heretical to call it an axiomatic *yukti* but that is what it is – each one of which is defined within a general logical-arithmetical (axiomatic) structure that naturally incorporates bivalence.[9]

It is not hard to imagine why logicians in India would have looked with suspicion at a proof such as this. Firstly, they were steadfast in keeping away from mathematical and even grammatical models of their logic(s) which might have lent queries about bivalence a sharper edge; while insisting on high standards for the admissibility of arguments, they always remained bound to meaning and language with their inherent lack of precision, dealing in propositions that tried to encompass everything under the sun and beyond, from the mundane – a standard syllogism goes: there is smoke on the hill, smoke is caused by fire

[9]It may be helpful to distinguish between *reductio* proofs in which negation of a putative proposition is shown to be in contradiction with itself (not self-consistent) and those in which it violates some consequence of the axioms governing a general construct in which the proposition is formulated (standard arithmetic in the present case: unique factorisation and rationality of $\sqrt{2}$ are incompatible). The irrationality proof of π outlined in Chapter 13.4 is more direct in this respect: it shows that an assumed finite continued fraction expansion of $\pi/4$ is not self-consistent and the only arithmetic it calls upon is division with remainder.

(illustrated by the example of a kitchen, logically superfluous; even a deductive inference, it seems, must be inductively reinforced), hence there is fire on the hill – to the supernatural as in the tetralemma: there is an afterlife, or not, or neither, or both. The essential point about bivalence, that **not not** is affirmation, $(S')' = S$, needs at the least a delimitation of the universe of propositions within which the negation of any given proposition is to be sought. Once that is done, there is no logical distinction between affirmative and negational statements: S and S' are symmetric with respect to the operation of **not**. The fact that all schools of Indian philosophy did make a distinction between affirmation and negation – some logicians even admitted the exclusion of the middle as a valid argument in negation but not in affirmation – is the clearest pointer to their lack of interest (to put it no more strongly) in this very powerful tool of demonstration. The bigger surprise is that no mathematician consciously broke through the barrier, even if examples of bivalent reasoning do occasionally appear, as though in moments of philosophical forgetfulness; Brahmagupta had no difficulty with the rule that positivity defines a bivalence on the set of all non-zero numbers: $(-)^2 = +$.

I conclude by noting that *reductio* methods are particularly well adapted for irrationality questions since the natural definition of an irrational number is negational, a number that cannot be expressed as a/b. (Attempts at an affirmative definition – what is it then? – had to wait till the second half of the 19th century). Indeed, all elementary proofs of irrationality are *reductio* proofs and it is not clear that even very sophisticated proofs are fully free from the taint of reasoning by contradiction in their total architecture. In principle, one can imagine that an assumed rationality of say π can be shown not to be self-consistent without using the knowledge that it is a real number; the continued fraction proof comes close to it. Did Nilakantha have some such method in mind when he made his conjecture? We shall probably never know.

15.3 Recursion, Descent and Induction

Just as conspicuous as the repudiation of proofs based on the excluded middle is the ubiquitous presence in Indian mathematics of recursive techniques in a variety of roles. The words 'recursive' and 'recursion' have acquired rather precise connotations in modern times, in logic, mathematics, computational theory, linguistics, and so on. But because the technique crops up in Indian mathematical (and nonmathematical) thinking in so many different guises – which are yet to recieve the analytic attention they deserve – it seems best to adopt, as I have done throughout this book, a somewhat loose characterisation: a recursive structure is a chain of steps, each of which produces an output for a suitable input, which can be used as the input in the next step. It has the merit of accommodating the diverse ways in which the basic recursive idea is put to profitable mathematical use, ranging over definitions, constructions and algorithms both exact and approximate, generation of infinite series for numbers

and functions, proofs, and so on. Given the amorphous form in which texts package and present their mathematics, it is also natural that the boundaries between these various subdivisions cannot always be sharply drawn. At any rate, having seen so many examples of the idea at work in the first three parts of this book, we can limit ourselves in this summary overview to some general observations, both of a historical and technical nature, together with references to some of the examples we have already met.

That recursive structures in nonmathematical contexts are met with very early in Indian culture has been known for some time now, thanks chiefly to the pioneering work of Frits Staal on the rule-based (syntactical) organisation of Vedic rituals and their accompanying chants (*mantra*s). The component parts of a ritual, some of them lasting many days, are sequenced in a pattern that is rigidly prescribed, and the total pattern includes repetitions, substitutions, interpositions, exchanges, etc. of its subunits for which the instructions are given in a formulaic way; they can be iterated indefinitely within the constraints of the means available and so can the chants ([St-RM]). In the *prātiśākhya* readings, which were composed at about the same time that the rituals were formalised and frozen, we can see a similar pattern of recursive reorganisation of an original text applied to its phonological subunits (Chapter 1.3). Staal's thesis is that the same recursive mindset inspired its extension to grammatical subunits, most clearly visible in nominal composition in which a compound of two words formed by applying one rule can be further combined with a third word, subject not necessarily to the same rule, endlessly in principle. That leads on straight to the naming of numbers, as we have seen. The number names of the *Ṛgveda* illustrate perfectly the potential of nominal composition for the creation of a rule-based recursive nomenclature for the infinitude of numbers (see Chapter 5): " ... there is no end to the names of numbers" (*Yuktibhāṣā*). That is characteristic of recursive constructions in general: one only needs a finite number of units (here, of language) and a finite number of rules for their manipulation to generate an infinite number of different, complex, expressions.[10]

Mathematically, the primordial recursive construct is the set of natural numbers, whether viewed as generated by the operation of succession or decimally through multiplication by 10, iterated as necessary, followed by addition. In the decimal viewpoint, names of numbers are faithful grammatical realisations of the more abstract arithmetical recursive structure of the numbers themselves. (It also leads, when algebraised as *Yuktibhāṣā* does, to a notion of a polynomial as a recursively constructed object but that is far into the future).

[10]It is also the moral of Patanjali's story about Indra learning grammar (Chapter 14.2). In the case of number names, the finite set of units comprises the names of atomic numbers and the rules are, ideally, three in number, implementing addition, multiplication (by powers of 10) and exponentiation (to powers of 10). Regrettably, the Vedic enumerators and all those who followed did not utilise the last operation (which I have therefore ignored in this book), and were obliged to invent new names for higher and higher powers of 10, infinite in number, causing themselves a lot of confusion thereby in the naming of powers of 10.

Historically, the interesting fact is that the earliest source for these different manifestations of recursion is the same. The early ('clan') Books of the *Ṛgveda* are already rich in decimal numbers (a fact which the philologists of Vedic language did not pay much attention to for a long time); likewise, general nominal composition (including in the making of number names) is a fully evolved part of the syntactical apparatus of the *Ṛgveda* (though the systematisation of formal rules had to wait for many centuries to elapse); and, similarly, though prescriptions for structuring the main Vedic rites came only around the 8th century BCE, about the same time as the earliest editions of the *Śulbasūtra*, the rites themselves have an older prehistory – the chants have many extracts from the *Ṛgveda* and parts of the Veda of rituals, *Yajurveda*, are probably even older as evidenced by its many examples of comparison counting (see Chapter 5.1). The Vedic civilisation was born with a dominant recursive gene and the question of which particular expression of it came first (and influenced the others) would seem to be a profitless line to pursue.

To the recursive construction of all decimal numbers, there is an inverse process, that of determining the decimal representation of any abstractly conceived (finite) number (the cardinality of a finite set). This too is recursive, the elementary step that is iterated being division (by 10) with remainder (Chapter 4.2). It is an example of finite recursion, the simplest such, as the iteration terminates when the quotient in a particular iterate becomes less than 10. A less elementary illustration of finite recursion is the *kuṭṭaka* solution of the linear Diophantine equation, based as it is, also, on division with remaider, the Euclidean algorithm (Chapter 7.2). It is also a typical example of the method called descent later in European number theory, the essence of which is the recursive reduction of a given problem to a simpler one in which relevant (positive integral) numerical coefficients are made smaller until, after a finite sequence of such steps, the problem becomes simple enough to be solved 'by inspection' or to be seen not to have a solution. (The determination of cardinality by the division algorithm may be considered to be the prototype of finite descent). The solution of the linear Diophantine equation is of the first type while European descent proofs are often of the second (non-existence) type. Interestingly from the Indian point of view, the latter type often involves a combination of recursion with *reductio ad absurdum*, as in the proof of the irrationality of π given in Chapter 13.4; no such proofs are to be expected in Indian texts.

Aryabhata's trigonometry invokes a variant of recursion which is also finite and which we may term cyclic: the output of some step in the chain coincides with the input/output of an earlier step. If no approximations are made in the initial values of the sine and the cosine (which is the input in the first step) or in the recursive solution of the difference equations, i.e., if we follow Nilakantha's treatment of the problem (Chapter 7.4), the periodicity of the sine will ensure that the output of the 96th step (for the canonical 24-fold division of the quadrant) will be the input of the first step or, by complemetarity of the sine and the cosine, that the output of the 24th step will be the input of the first step in the computation of the cosine. (Aryabhata's own rule which

is subject to approximations in the first two differences exhibits cyclicity only approximately).

In the Nila school, the philosophy was extended to true functions: the determination of functions satisfying some equation, algebraic or differential, given their value at a point (initial value problems, the computation of the sine table being a discrete example of the same type) by a process of recursive refining (*saṃskāram* in *Yuktibhāṣā*'s terminology). Calculus is about functions even if they originate in geometry; so this was a natural extension of an old technique which, in the process, acquired a depth and versatility unseen before. Little needs to be added to the many examples of this type of problem we have seen in Chapters 10 to 13 (and the general remarks of Chapter 10.4) except to recapituale briefly certain features they have in common. In its simplest form, if f is the function to be determined, and x_0 the point at which it is known, the basic idea is to start with the identity $f(x) = f(x_0)+(f(x)-f(x_0))$ and use the equation satisfied by f to recast the second term on the right (the correction) in a form into which the same equation can be substituted. The sine series which starts out with the sine difference formula (equivalently, the second order *saṃkalitam*/integral equation), together with the initial values $\sin 0 = 0, \cos 0 = 1$, is the most sophisticated example. It is a little more elaborate than the description above because the difference equation involves the cosine. There are other variations possible: the function may well be 'known' and what is sought may be a reexpression convenient for the carrying out of other operations, e.g., the geometric series for $(1 + t^2)^{-1}$ to facilitate integration in the derivation of the π series (Chapter 11.3). In yet another variation, when it is not possible or practical to express $f(x) - f(x_0)$ in a form convenient for iteration, a clever approximation may have to be resorted to; the hope, not always justified (e.g., the interpolation series, Chapter 10.3-4), is that the error so introduced will tend to vanish with increasing order of iteration. The prototype of such an approximation scheme is the Bakhshali square root formula (Chapters 6.5, 10.4; perhaps also the *Śulbasūtra* $\sqrt{2}$) iterated indefinitely – the initial value is that at the nearest perfect square and the difference equation which involves a square root is linearised by means of Heron's formula – and the most demanding is the estimation of the truncation error in the π series (Chapter 13.3).

As seen in these examples, iterated refining, though often truncated in calculations of approximate corrections, is non-terminating by construction. Consequently, whenever an exact answer is sought, the convergence of the formal series generated has to be independently established. Both Jyeshthadeva and Nilakantha worried about this question: the former in stating that the correct variable to expand the angle in when it lies in the second octant, $\pi/4 < \theta < \pi/2$, is $\cot\theta$ rather than $\tan\theta$, to ensure that the variable is bounded in absolute value by 1 (Chapter 11.5) and the latter in using the geometric series to illustrate the necessity of thinking of the sum of a power series as a limit (Chapter 10.4). As noted in the same section, the geometric series also provides a simple illustration of the limitations of formal refining *ad infinitum*; it converges only if $t^2 < 1$.

Qualitatively different from all these technical applications is the use of the idea of recursion as a metamathematical device, namely the proof technique now called mathematical induction. In essence, it consists in establishing the truth of a set of propositions which are in one-one correspondence with natural numbers by exploiting the recursive structure of the latter. Its axiomatic foundation is the identification of the successor of *any* non-negative integer with the result of adding 1 to it, which is one way (the *Yuktibhāṣā* way) of thinking of the induction axiom of Peano. The first ever systematic exposition of the method is in *Yuktibhāṣā* in connection with the integration of positive powers (see Chapter 11.3-4) and the care Jyeshthadeva takes in describing the proof is a measure of his recognition of its absolute novelty.[11] So is his careful identification, within the decimal paradigm, of the operations of addition and succession (see the quotation in Chapter 4.2).

If the crowning achievement of the recursive mode of thinking is mathematical induction, it is a pleasing thought that it has led us back to where it all began, to the very first of all recursively defined mathematical objects, the set of natural numbers.

[11]It is sometimes said that Euclid's proof of the infinitude of primes is inductive. That is about as true as the opinion that Greek circle geometry is trigonometry. (There are inductive proofs of the proposition but they are modern). It has also been stated by some historians that Aryabhata's *kuṭṭaka* and Jayadeva's *cakravāla* are early examples of inductive algorithms. Being integral solutions of the relevant algebraic equations, constructed by descent and hence of the finite recursive type, they have nothing logically to do with the infinitude of natural numbers or the induction axiom.

16

Upasaṃhāra

The title of this short final chapter has many meanings in Sanskrit and Malayalam: conclusion, summing up, drawing together, a collection, etc. What is collected here are a few general remarks, all in one place, about the state of mathematics in India – as Nilakantha in the wisdom of his years might have envisioned it, say – at the apogee of its trajectory. The word *upasaṃhāra* also means destruction, the end.

16.1 Towards Modernity

Of the mathematical peaks the Nila school scaled, none is taller than the invention of calculus. In our admiration for the originality and technical virtuosity of that accomplishment, it is easy to overlook the many other advances, both conceptual and methodological, that accompanied it, some of them an integral part of the calculus itself, others related to it only indirectly. Pervading all this however, and even more significant historically, is a new perception of mathematics as a self-sufficient and self-defined discipline. The very strong astronomical bias it had acquired with and after Aryabhata had little actual influence on Madhava's – though he was a (reluctant) practising astronomer as well – infinitesimalisation of trigonometry and on the technical tools he shaped in order to realise his goals. The astronomy of the time (or even of today) did not require an exact value for π or a knowledge of $\sin\theta$ for *every* value of θ. It is an incongruous experience to read in the first line of *Yuktibhāṣā* that it is an account of the mathematics required in Nilakantha's astronomical text *Tantrasaṃgraha* and then to see chapters 6 and 7 – well more than half the book – unfold before us a mathematics which is of little relevance to astronomical applications; having given his tribute to an old tradition, Jyeshthadeva seems to have felt free to let the mathematical imperative take over. Mathematics has become its own motive force.

There were of course earlier people who pursued interesting mathematics for its own sake, notably Brahmagupta in his work on number theory and cyclic

P. P. Divakaran, *The Mathematics of India*, Sources and Studies in the History of Mathematics and Physical Sciences, https://doi.org/10.1007/978-981-13-1774-3_16

geometry. Neither the setting up of the quadratic Diophantine equation nor the focus on the transformations of its space of solutions as encoded in the operation of *bhāvanā* really derives from any problem in astronomy though Brahmagupta pretends otherwise (much like Jyeshthadeva did later about the purpose of his book). Similarly, the resort to symmetry transformations (transpositions of sides) to prove surprisingly novel theorems on cyclic quadrilaterals culminating, ultimately, in the area theorem and the diagonal theorem, owes little to Aryabhatan trigonometry or astronomy. The trend continued with Brahmagupta's true apostle Narayana and, as far as cyclic geometry is concerned, on to *Yuktibhāṣā*. It is not without reason that these topics and their animating ideas and techniques, especially the iterated application of *bhāvanā* to generate an infinite number of (and, after *cakravāla*, all) solutions of the quadratic Diophantine equation, have found strong resonance among mathematicians of the modern period; they are very modern. The use of permutation symmetry in the theory of cyclic quadrilaterals is no less modern in spirit, though hardly commented on by historians. These developments together open for us a window to Brahmagupta's structural approach to purely mathematical questions: behind them both lurk the embryonic notion of a group of transformations.

In the work of the Nila school, these early intimations of a structural mindset became an intrinsic part of its perception of the mathematical universe. We cannot doubt that the catalyst of the transformation was Madhava's genius for it is central to many of the insights and results attributed to him by his followers. The modernising trend covered virtually all subdisciplines but is, naturally, most pronounced in the analytical framework without which problems in calculus could not be properly posed. And, within that, the most fundamental notion is that of a (geometrically defined) function as the assignment of a (numerical) value to *every* point in the natural domain of a variable: $\tan\theta \to \theta$ for $0 \leq \tan\theta \leq 1$; $\theta \to \sin\theta$, first for $0 \leq \theta \leq \pi/2$ and then extended to the whole circle (there is a subsection in *Yuktibhāṣā* on the definition of the sine and the cosine in all four quadrants). For the modern reader for whom the idea of a (real) function of a (real) variable holds no mysteries, it may take an effort to appreciate what a great change in the ethos of Indian geometry it represents. We only have to reread Bhaskara II's accounts of the sine table and the sphere formulae – not to mention Bhaskara I's misunderstanding of $\pi/48$ as the smallest angle that can be – to realise how elusive the idea of a continuously varying function had been (as it was in other mathematical cultures).

Understanding what a function is was critical for the next step, that of defining its infinitesimal variation or differential. The Nila prescription is direct: to integrate a function $f(x)$ over an interval $[a, b]$, divide the interval into n equal parts at points x_i, sum the product $(x_{i+1} - x_i)f(x_i)$ and let n grow without bound. This is basically the definition of the Riemann integral. The originality is in the 'division by infinity'; working with quantities which are infinite but not quite infinite – rather than directly with quantities which are zero but not quite zero as for Newton – it is immune to the kind of metaphysical criticisms that infinitesimal calculus faced later in Europe.

The concept of a general function or map appears also in a non-geometric avatar in the definition of the remainder when the π series is terminated after an *arbitrary* number of terms, i.e., as a function on the set of positive odd integers, and its determination as the solution of the resulting functional equation, together with an absolutely clear-cut warning about the need to respect the dependence of the function on the variable (Chapter 13.1).

Among the other steps towards an ever more specifically mathematical culture, I recall again and without elaboration the following: the first recognition of infinite series (including power series in one variable) as legitimate mathematical entities, along with (necessarily rudimentary) considerations about their convergence; the assured handling of infinity itself, free from the constraints of nominalism; a clear understanding of what a variable is – it is not just the unknown, the *yāvat tāvat*, of earlier authors – and an abstract algebraic definition of polynomials (and rational functions) in one variable and the operations that can be performed on them (and, even more unexpectedly, the recognition itself that they needed to be defined).

The opening up of these fresh avenues went hand in hand with an emphasis on rigour and clarity in the foundations as well, a new set of rules of the road as it were. Proofs are well organised, explicit and complete. Where time-tested approaches could no longer be generalised to meet the needs of the new mathematics, entirely original proof methods were invented, mathematical induction having pride of place among them. And to lend logical support, old and settled issues were revisited: decimal enumeration and the arithmetic of addition were recast so as to bring to the fore the propery of succession of the natural numbers. It is not quite Peano axiomatisation, but it is, nevertheless, remarkable for its recognition that mathematical induction rests on the identification of the successor with the result of adding 1.

It is difficult to say how much of this concern with method goes back to Madhava himself and how much is part of the Nilakantha-Jyeshthadeva project of freeing mathematics from the often unrigorous habits (going by the writings) of the past. But there can be no doubt that the brilliant 'hard' mathematics for which it served as the foundation was almost all Madhava's. There is an embarrassment of riches to choose from, but the outstanding example has to be the derivation of the sine series which, in the modern vocabulary of calculus, can be summarised in the following sequence of steps: convert the derivatives of sine and cosine (a pair of coupled first order equations) into a second order differential equation for the same functions; turn it into a (Volterra) integral equation by formal integration; and solve by iteration. These are techniques which belong to the 18th and 19th centuries in Europe. Similarly, we have to wait till the middle of the 18th century to see matched, most notably in the work of Euler, the kind of algebraic and analytical power that is displayed in Madhava's manipulation of the π series to make it more amenable.

A final word (for the last time!) about the Nila school's favourite technical tool, iterated refining or *saṃskāram*. Stripped of practical issues like the approximations employed, the elementary step that is iterated is substitution.

In the modern way of thinking about it, a substitution is just a composition of maps (that can, in terms of the rules that define the structure of the sets involved, be composed) and since compatible maps can be composed *ad infinitum*, a *saṃskāram* is a natural procedure in the framework of sets with structure. The best illustration is, again, provided by the integral equation for the sine which I restate:

$$\sin\theta = \theta - \int_0^\theta d\phi \int_0^\phi \sin\chi d\chi.$$

This is a condition that can be thought to define the sine as a map from the circle to the interval $[-1, 1]$. For the Nila mathematicians, it also determined the sine as a power series by the simple substitution of $\sin\chi$ by its definition, and so on indefinitely. It is a purely algebraic process as is clear from its use in the discrete version of the same problem (see Chapter 12.2) – not to mention its remote ancestor, the Bakhshali square root – and is an excellent and sophisticated example of what Mumford calls the reification of a mathematical problem, though it is done without symbols.[1]

The remarks above are not about establishing priority. The age of mathematical enlightenment that began in Europe with the Cartesian revolution has grown from strength to strength and continues to flourish as never before. What it has created is a universal model of mathematics whose values and principles are now unquestioned by the mathematical community as a whole (despite sporadic expressions of reservations about foundations). Certain components of it may have come from Babylonia or Greece or India, but it is by the standards of this new paradigm that we must judge the worth of earlier contributions to its making. By that yardstick, the new directions the mathematics created and nurtured in the Nila basin took pointed unerringly to the future; it was discovering for itself the very same universal elements – in the sense Proclus gave to the title of Euclid's book – that Europe later built upon and made into the magnificent edifice mathematics is today. In the process, Indian mathematics became less Indian so to say. Perhaps it is a truism that every great advance in mathematics, whether in making theories or solving problems, is a step towards its universal core. Perhaps, at the highest level, there is no difference between the two.

16.2 Cross-cultural Currents?

In the course of this book we have come across instances of mathematical (and related) ideas travelling across the vast and intractable distances that sepa-

[1]David Mumford, "The Invention of Algebra as Reification", lecture at the International Conference on Mathematics in Ancient Times (2010) in Kozhikode (Calicut) (available on the web at http://www.dam.brown.edu/people/mumford/beyond/papers/2003c–ICMATtalk.pdf). For Mumford, algebra is substitutability.

rated ancient cultures, to become agents of intellectual renewal in their new homes: the journey of the Ptolemaic planetary model from Alexandria to the India of the Siddhanta period and, reciprocally, of the astronomy and mathematics of Brahmagupta and Aryabhata to Islamic Mesopotamia and China, for example. We have also had occasion to consider at some length suggestions of trans-cultural transmissions which, on examination, turned out to be unsupported and unviable speculation: Indus-Babylonian-Vedic-Greek geometric interactions (Chapter 3.4), trigonometry from the Mediterranean to Aryabhatan India (Chapter 7.4) and, in the other direction, the fundamental idea of calculus from Kerala to Europe (Chapter 12.4). There are lessons to be drawn from these examples (and others like them) among which none is more important for the historian than a very general one: settling questions of likely transmission is a complex undertaking involving (to list only the most obvious issues) chronological precedence; physical evidence of contact, preferably documentary, and a plausible mechanism of transmission; and, demanding but potentially most rewarding, a rigorous assessment of the imprint of the intruding mathematics on the host.

Temporal priority is generally reasonably easy to establish but has not been always above controversy – recall the efforts of British colonial Indianists to push forward the earliest *Śulbasūtra* to a post-Euclid date, in the face of linguistic evidence to the contrary. But, since everyone knows (except, going by what they wrote, the orientalist historians themselves) that priority is only a necessary but not sufficient criterion for the identification of what is cause and what is effect, it is often sought to be reinforced in more recent times by what I have earlier called mathematical snippets: Ptolemy not only lived before Aryabhata but also made a table of chords somewhat similar to Aryabhata's table of sines; the same for Hipparchus who, if we read the tea leaves correctly, *may* have used a value of 3438 minutes for the radius of a circle. As we saw in Chapter 7.4, such wishful appropriations do not pass any sort of an examination that is not totally superficial. (They do get repeated though).

Towards the middle of the last century a new twist was given to the primacy of precedence by the mathematician van der Waerden:[2] the great discoveries of mathematics (among which he counts the diagonal theorem and the solution of the quadratic Diophantine equation) are so singular that it is inconceivable for more than one (presumably, singularly great) mind to have made them. Consequently, a subsequent rediscovery of a significant result must be attributed to transmission, never mind by what means. This was a true gift to lazy historians: priority is all; once accepted, it relieves them of the burden of looking for anything beyond dating techniques. And if there is no evidence of priority, it can always be endowed: van der Waerden himself "supposed" (his own word) that the quadratic Diophantine equation was posed and solved by the Pythagoreans and, further, that the solution was "copied without proof" by

[2]B. L. van der Waerden, *Geometry and Algebra in Ancient Civilizations*, Springer, Berlin (1983) (as well as much other writing of his) has an account of his views on the subject.

Brahmagupta. Since there is no Greek text that Brahmagupta can be checked against, there is also no question of subjecting van der Waerden's dictum ('the dictum') to any sort of test; it is history by fiat.

This would not be a serious issue if van der Waerden's was a lone voice that fell on deaf ears. Astonishingly, it is not; the dictum, though not always consciously acknowledged, has a following among a serious section of contemporary professional historians, as we have seen several times. If there is an explanation for the continued popularity of this misguided conviction that had its origin in Europe's imperial-colonial past, it has yet to be openly put forward. We can only amuse ourselves by imagining how the dictum might respond to other, unquestionably securely dated, instances of relative antecedence, for instance the dilemma posed by Madhava's prior derivation of the series associated with the names of Newton and Leibniz, not to speak of his discovery of the fundamental principles of infinitesimal calculus. Unfortunately for us, van der Waerden died before the work of the Nila school came to be known to the world (outside a narrow circle and despite Sarasvati Amma's splendid book and the papers of Rajagopal); so we do not know how he might have come to terms with it. His intellectual followers of today are happy to say that the series came out of Indian mathematicians' supposed computational wizardry combined with some sort of 'calculus' within quotes. The irony is that it is precisely the same dictum that is the most forceful argument in the case made by the protagonists of the hypothesis of a transmission of infinitesimal calculus from Kerala to Europe. And it fails the same tests; there is no evidence of any mathematical contact – despite the constant comings and goings of Portuguese mariners/merchants to and from Kerala during the 16th and the first half of the 17th century – nor any great similarity in the way the universal central concept of calculus, local linearisation and its inverse, is implemented; the invariant calculus gene mutated differently in India and in Europe.

The moral we can draw from these expressions of what is no more than opinion is that, where mutual influences are concerned, an objective judgement can only be pronounced after a careful analysis and comparison of the internal mathematical evidence. A useful model for such an exercise is the discipline of comparative linguistics. Linguists have known for a very long time that, in order to test the hypothesis of a common ancestor language, say for Sanskrit and Latin (the putative Indo-European), it is not enough to find a few common words (with, roughly, the same pronunciation and the same meaning; the equivalent of our mathematical snippets) or even many such, as is actually the case. Far more weighty evidence is provided by commonality in structural features like grammar, syntax, usages allowed and forbidden, the taxonomy of words and phrases and so on. Already at the end of the 18th century, William Jones was speaking of the strong affinity between Sanskrit and Greek and Latin "both in the roots of verbs and in the forms of grammar". It is only after the common structural frame is reliably identified that one can begin to look for differentiators, the culture-specific mutations of the invariant linguistic gene.

(The analogy between the evolution of biological organisms and of languages has turned out recently to be a useful one for linguists).

The difference between language and mathematics is that, at a fundamental level, there is only one mathematics, absolute and immutable. This is not to be taken necessarily as a statement of philosophical position but as an empirical observation: the diagonal theorem is the same for Babylonians, Vedic Indians and Pythagoreans (and in today's classrooms); likewise, the solution of the quadratic Diophantine equation is the same for Brahmagupta-Jayadeva and for Euler-Lagrange (though their methods may appear to be quite different at first sight). The linguistic analogy would be if every language in the world were descended from Indo-European or every human brain wired to process the structure of Indo-European alone rather than language generically. That makes the study of the comparative history of mathematics superficially easier and at the same time more difficult: easier because the quarry is the same and, so, easy to identify; and harder because, precisely for that reason, we cannot stop there. To come back to where we started from, finding a common mathematical truth in two apparently isolated civilisations proves nothing about a possible interaction between them.

As a matter of operational discipline (Nilakantha would have called it *vyavahāra*), it seems then best to accept that mathematical truths are absolute and immutable, as likely to be discovered at a given time and place as at any other, given propitious circumstances; we might even take it as the only workable definition of their universal identity and recognisability. Without such a principle to guide us, comparative history of mathematics will either lapse into absurdities like van der Waerden's dictum or cease altogether to be a useful discipline. Or – for those who do not care for the whiff of Platonism such a stand implies – we can switch to Aryabhata's very Indian metaphor: to bring up the best of mathematical gems from the ocean of true and false knowledge, all one needs is an intelligent mind for a boat; it does not matter where the boat is made.

16.3 Journey's End

This book is now at the end of its long voyage. The overriding impression one is left with is of an outstandingly gifted community of mathematicians, passing the flame from generation to generation, never large in numbers, respected but not very influential outside its own narrow circle. There were good times, times when they scaled great heights,[3] led there by some of the finest mathematical

[3]Contrast the way mathematics renewed itself periodically to face the fresh challenges thrown up by its own successes with the situation in the other science with which it is commonly paired, *āyurveda*, the science of life. There has been no fundamental changes made to the doctrine of *āyurveda* after the great triumvirate of antiquity, Caraka, Suśruta and Vāgbhaṭa (the last probably a contemporary of Aryabhata), neither in the understanding of the biological processes in the human body nor in the theory and practice of diagnosis and treatment.

minds of all time, and there were dark times through which the lamp was kept burning, flickering but not extinguished. Through it all, they looked to their own masters from the past, the *pūrvācārya*, for inspiration; the mathematics remained self-sufficient, one might say self-contented, even insular. How else to explain that, at a time when northwestern India was the epicentre of Indian intellectual life and was awash in Hellenic culture in all its aspects, the only scientific idea felt to be worthy of absorption was that of planetary epicycles? Why were Euclid's methods and theorems in geometry ignored or shunned? Whatever the reason, the loss was India's. Unlike the Greeks, India never thought that mathematics could be used with profit in areas of science other than descriptive astronomy, did not produce an Archimedes of its own; did not even seem to have heard of him.

One can argue that the secret of the longevity of the mathematical life of India is this very conservatism of those who kept the faith. Wars and invasions came their way, but they were willing travellers, either to escape misfortune or in search of green pastures and generous patrons. And no pasture was greener than a little patch of the Zamorins' domain and it is there that the story came to an end in a final display of brilliance.

The end when it came was sudden. Nilakantha could not have imagined when he set down the epistemic foundations of good mathematical and astronomical practice for future generations – and looking at the work of his pupil Jyeshthadeva – that there were not going to be too many future generations; that, within less than a century, there would be no one left to build on the legacy Madhava bequeathed and he himself helped nurture. Achyuta was for all practical purposes the last of the line and he was a part-time mathematician. Works of a derivative nature continued to appear at infrequent intervals, as they did till very recently, but there was not much that was genuinely creative in them. Some of those who might, in good times, have become astronomers turned to the profession of horological astrology for which there was – and still is – a market; Alattiyur, once home to Parameshvara and Nilakantha, has a thriving and long-established clan of astrologers who count among their clients and patrons the powerful of the land. The last of the traditional colleges, the *gurukulam*s ('lineage of teachers') training the young in Sanskrit and the *śāstra*, had shut their doors by the middle of the 20th century.

The death of mathematics on the Nila was also its death in India. Northern India never really recovered from the atrocities of Ghazni and Ghori. A short-lived renaissance took place in Benares when the Daivajña family migrated there from Maharashtra towards the end of the 16th century, at about the time the decline of the Nila school was gaining speed. And, even while it lasted, the renewal produced not much that was original, contenting itself with elucidations mostly of the work of their illustrious compatriot Bhaskara II. After that there is only silence.

The immediate trigger for the eventual disappearance of mathematics from the Nila basin was yet another invasion, that of the Portuguese

under Vasco da Gama in 1498, two years before Nilakantha completed *Tantrasaṃgraha*. They too came for the loot, not of the temple treasures of Malabar but of its green wealth, the spices whose trade had made Muziris, a hundred kilometers to the south, a renowned emporium in the antique world at the turn of the common era. Trade was the pretext but it was trade under the gun. The Zamorins had made Calicut a new Muziris, more famous and more prosperous than the old, and it was control of the spice trade through military dominance that the Portuguese sought. The two or three decades that followed (Vasco died in Cochin in 1524) saw a great deal of fighting on sea and on the shore between the mouth of the Nila and Calicut, the very land that sustained the temple-villages and their mathematical life. This was the time when Nilakantha would have been thinking his way towards the insights of his last great works; Jyeshthadeva would have been in the prime of his life, his lone masterpiece gestating in his mind, waiting to be written.

The confrontation with the Portuguese had its ups and downs but the Zamorins were the real losers: it emptied the treasury of Calicut. The wealth of the great temples was depleted and Calicut's fame as a centre of institutional learning, its academies and its famed annual assembly of scholar-rivals were all washed away in the flood of violence and bloodshed. The glory never returned. The Portuguese were eventually driven off the Kerala coast, to be succeeded by the Dutch and they too contributed their bit to the vandalism by burning down a Jesuit seminary and library near Cochin, reputed to have held an extensive collection of Indian manuscripts. They in turn were pushed out by the English empire builders in the guise of the East India Company, which had by then (early to mid 1700s) planted the seeds of the future empire virtually all over India.

The world had come to the doorstep of the Nilakanthas and the Jyeshthadevas and their own world had been changed for ever in consequence. Even if a few hardy souls among their disciples were prepared, like their forefathers, to set off on another long journey from their home of half a millennium, where would they have found a sympathetic king or lord to give them honour and a decent life?

In any case, colonial visitations were perhaps only the trigger, they did not bring the curtain down on the Nila phenomenon unaided. By the 15th century, the depressingly short-sighted social mores of the Namputiris had begun to catch up with them. The custom of restricting to the eldest son alone the right to marry within the Brahmin fold and beget Brahmin offspring had led to a precipitous fall in the share of Namputiris in the population. As a means of keeping academic traditions alive, sons of Namputiri fathers and non-Namputiri mothers were permitted to get an education in the sciences and the art forms (but not in the sacred Vedic lore); Shankara and Achyuta, as we know from their caste names, are examples. But they came too late and they were too few. The Brahmins had seen to it that the vast majority of males in Kerala (women were in any case kept illiterate), in fact all who lacked a Namputiri Y chromosome, had no right to an education of any sort.

And that really is the end of our story. When the British established their own schools and colleges in 19th century India, the mathematics that was taught owed nothing to the long Indian tradition and to the labours of a Colebrooke or a Playfair or a Whish to bring it to light. It was a fresh beginning and it was not long before Ramanujan brought mathematics at the highest level, the new mathematics, vibrantly to life again.

Bibliography

(A selection of references which I have found useful. Not all are referenced in the text.)

Sources with translations and/or annotations.

(No reference is given for the *Ṛgveda*; there are a large number of translations and interpretations in a large number of languages. Precise references to chapter, poem and verse for questions related to number names can be found in Bhagyashree Bavare and P. P. Divakaran, "Genesis and Early Evolution of Decimal Enumeration: Evidence from Number Names in *Ṛgveda*", Ind. J. Hist. Sci. vol. **48.4** (2013) p. 535).

[AB-C] The *Āryabhaṭīya* translated into English with commentary by Walter Eugene Clark, University of Chicago Press, Chicago (1930).

[AB-S] *Āryabhaṭīya* of Āryabhaṭa, critically edited with English translation, notes, etc. by Kripa Shankar Shukla in collaboration with K. V. Sarma, Indian National Science Academy, New Delhi (1976).

[BM-H] The Bakhshali Manuscript, with translation and commentary by Takao Hayashi, Egbert Forsten, Groningen (1995).

[GK-S] *Gaṇitakaumudī* of Nārāyaṇa Paṇḍita, translated with notes by Parmanand Singh, published periodically in instalments in Gaṇita Bhāratī, (1998 - 2002).

[Ne-W] The Mathematical Papers of Isaac Newton (vols. I - VIII), edited with notes and translations from the Latin, where necessary, by D. T. Whiteside, The University Press, Cambridge (1967 onwards).

[PS-K] *Pañcasiddhāntikā* of Varāhamihira with translation and notes by T. S. Kuppanna Sastry, P.P.S.T. Foundation, Madras (1993).

[SiSi-P] L'Œuvre mathématique et astronomique de Bhaskaracarya: Le Siddhantasiromani I-II, with translation and commentary in French by François Patte, Librairie Droz, Paris (2004).

P. P. Divakaran, *The Mathematics of India*, Sources and Studies in the History of Mathematics and Physical Sciences, https://doi.org/10.1007/978-981-13-1774-3

[SuSi-B] The *Sûrya Siddhânta*, translation with notes by Ebenezer Burgess, Motilal Banarsidass, Delhi (1989, reprint).

[SuSu-SB] The *Śulbasūtras* of Baudhāyana, Āpastamba, Kātyāyana and Mānava, with translations and commentary by S. N. Sen and A. K. Bag, Indian National Science Academy, New Delhi (1983).

[TaiSam-K] The Veda of the Black Yajus School entitled *Taittirīyasaṃhitā*, translated by A. B. Keith, Harvard University Press, Cambridge (1914).

[TaSa-RS] *Tantrasaṅgraha* of Nīlakaṇṭha Somayājī, with translation and commentary by K. Ramasubramanian and M. S. Sriram, Hindustan Book Agency, New Delhi (2011).

[YB-S] *Gaṇita-Yukti-Bhāṣā* of Jyeṣṭhadeva, vol.I (Mathematics), Malayalam text with English translation by K. V. Sarma (and explanatory notes by K. Ramasubramanian, M. D. Srinivas and M. S. Sriram), Hindustan Book Agency, New Delhi (2008).

[YB-TA] *Yuktibhāṣā* (Part I - General Mathematics), Malayalam text with *vyākhyā* in Malayalam by Ramavarma (Maru) Tampuran and A. R. Akhileshvara Ayyar, Mangalodayam Ltd., Trichur (1948).

To these has to be added Colebrooke's translation with notes of *Līlāvatī* and *Bījagaṇita* and chapters XII and XVIII of *Brāhmasphuṭasiddhānta*:

[Co] Henry Thomas Colebrooke, *Algebra with Arithmetic and mensuration, from the Sanscrit of Brahmegupta and Bháscara*, Sharada Publishing House, Delhi (2005, reprint).

Sources in the original (without translation or detailed commentary).

[ABB] *Āryabhaṭīya* with the *bhāṣya* of Nilakantha (on the *Gaṇita*, *Kālakriyā* and *Gola* chapters); Parts I and II, K. Sambasiva Sastri (ed.), Government of Travancore, Trivandrum (1930, 1931); Part III, Suranad Kunjan Pillai (ed.), University of Kerala, Trivandrum (1957).

[JM] *Jyotirmīmāṃsā* of Nilakantha, K. V. Sarma (ed.), Vishveshvaranand Institute, Hoshiarpur (1977).

[SD] *Siddhāntadarpaṇa* of Nilakantha with his own analytic commentary, K. V. Sarma (ed.), Vishveshvaranand Institute, Hoshiarpur (1976).

[TaSa-S] *Yuktidīpikā* of Shankara is part of *Tantrasaṃgraha* of Nilakantha (with *Yuktidīpikā* and *Laghuvivṛti*), K. V. Sarma (ed.), Vishveshvaranand Institute, Hoshiarpur (1977).

History of mathematics and related subjects

Books

[DS] Bibhutibhushan Datta and Avadhesh Narayan Singh, *History of Hindu Mathematics* vols. I and II, Bharatiya Kala Prakashan, Delhi (2004, reprint).

[If] Georges Ifrah, *Histoire universelle des chiffres*, Robert Laffont, Paris (1994).

[Jo] George Gheverghese Joseph *The Crest of the Peacock* (third edition), Princeton University Press, Princeton and Oxford (2011).

[Pl] Kim Plofker, *Mathematics in India*, Princeton University Press, Princeton and Oxford (2009).

[SA] T. A. Sarasvati Amma, *Geometry in Ancient and Medieval India*, Motilal Banarsidass, Delhi (1999).

[Sa-HKSHA] K. V. Sarma, *History of the Kerala School of Hindu Astronomy*, Vishveshvaranand Institute, Hoshiarpur (1972).

[St-RM] Frits Staal, *Ritual and Mantras: Rules Without Meaning*, Motilal Banarsidass, Delhi (1996, Indian edition).

[St-Veda] Frits Staal, *Discovering the Vedas. Origins, Mantras, Rituals, Insights*, Penguin Books, New Delhi (2008).

[Th-EI] Romila Thapar, *Early India from the Origins to AD 1300*, Penguin Books India, New Delhi (2003).

[Th-CP] Romila Thapar, *Cultural Pasts. Essays in Early Indian History*, Oxford University Press, New Delhi (2000).

Collections

[Agni] Frits Staal (ed.) *Agni. The Vedic Ritual of the Fire Altar* vol. II, Motilal Banarsidass, Delhi (2001, Indian edition).

[CHIM] Gérard G. Emch, R. Sridharan and M. D. Srinivas (eds.) *Contributions to the History of Indian Mathematics*, Hindustan Book Agency, New Delhi (2005).

[MEMCII] Victor J. Katz (ed.) *The Mathematics of Egypt, Mesopotamia, China, India and Islam. A sourcebook*, Princeton University Press, Princeton and Oxford (2007).

[SHIM] C. S. Seshadri (ed.) *Studies in the History of Indian Mathematics*, Hindustan Book Agency, New Delhi (2010).

[Sunya] A. K. Bag and S. R. Sarma (eds.) *The Concept of Śūnya*, Indira Gandhi National Centre for the Arts, New Delhi (2003).

Finally, there is the invaluable web-based archive of J. J. O'Connor and E. F. Robertson, unusually good on Indian mathematics. The master URL is

http://www-history.mcs.st-andrews.ac.uk/index.html

with links leading to various sub-indices.

Index

P. P. Divakaran, *The Mathematics of India*, Sources and Studies in the History of Mathematics and Physical Sciences, https://doi.org/10.1007/978-981-13-1774-3